PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, C12Q 1/68, G01N 33/68, C07K 16/18

A1

(11) International Publication Number: **WO 99/27092**

(43) International Publication Date: 3 June 1999 (03.06.99)

(21) International Application Number: PCT/GB98/03486

(22) International Filing Date: 23 November 1998 (23.11.98)

(30) Priority Data:
9724829.8 21 November 1997 (21.11.97) GB

(71) Applicant *(for all designated States except US)*: PROLIFIX LIMITED [GB/GB]; 91 Milton Park, Abingdon, Oxfordshire OX14 4RY (GB).

(71)(72) Applicants and Inventors: MÜLLER, Rolf [DE/DE]; Poitiers-Strasse 8, D-35037 Marburg (DE). LIU, Ningshu [CN/DE]; An der Haustatt 4, D-35037 Marburg (DE).

(74) Agents: BRASNETT, Adrian, H. et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: PURIFIED TRANSCRIPTION FACTOR CDF-1 AND ITS USE

(57) Abstract

The invention provides a method for obtaining the transcription factor CDF-1 in purified form, as well as assays which related to the modulation of CDF-1 repression of transcription induced by viral oncogenes, particularly HPV E7. Provided is an assay for a modulator of a viral replication, wherein said replication is facilitated by suppression of cdc25 promoter repression by CDF-1 due to the action of a viral protein, which assay comprises: a) bringing said viral protein into contact with one or both of: (i) a CDF-1 protein; and (ii) a DNA sequence comprising a CDF-1 binding site, in the presence of a putative modulator under conditions wherein in the absence of modulator the viral protein is able to induce inhibition of CDF-1 binding to DNA; and b) measuring the degree of modulation of binding of CDF-1 to DNA.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

PURIFIED TRANSCRIPTION FACTOR CDF-1 AND ITS USE

The present invention relates to the transcription factor CDF-1 and its use in assays for modulators of viral replication.

Background to the invention.

Transcriptional repression has turned out to be a major mechanism of gene regulation during the cell cycle (Ref.11, Ref.28, Ref.51). A group of genes has been identified whose transcription is blocked early during the cell cycle by complexes consisting of the transcription factor E2F (Ref.19) and pocket proteins of the pRb family (Ref.28, Ref.45, Ref.44), including B-myb (Ref.20), E2F-1 (Ref.14, Ref.17), p107 (Ref.50) and orc-1 (Ref.29). All these genes become derepressed prior to S-phase entry. A second group of genes which is expressed later during the cell cycle and seems to be controlled by a different mechanism of transcriptional repression has also been identified. When the cdc25C promoter, which is up-regulated in late S/G_2, was studied by *in vivo* footprinting and mutational analysis, a novel repressor element, the 'cell cycle dependent element' (CDE), was identified (Ref.24). The CDE is occupied in G_0-G_1 and its occupation is lost in G_2, when cdc25C is expressed. That CDE mediated repression plays a role in regulating other promoters as well was shown by the presence of functional CDEs in the cyclin A and cdc2 promoters which reach their maximum expression in S/G_2 (Ref.54, Ref.15).

These studies also led to the discovery of an additional element contiguous with the CDE, which is identical in all three promoters. This element was termed 'cell cycle genes homology region' (CHR) (Ref.54). Mutation of either the CDE or the CHR in the cdc25C, cdc2 or cyclin A promoter largely abolishes repression in G_0. These functional data were supported by the demonstration of G_0-G_1-specific protein binding to both the CDE and CHR in genomic footprinting. Interestingly, the CDE is contacted in the major groove of the DNA while the binding to the CHR occurs in the minor groove (Ref.54).

In summary, the CDE and CHR are bound by a transcriptional repressor in G_0/G_1 which is released in S/G_2 (Ref.24, Ref.54). The CDE/CHR-interacting factor has been identified and termed CDF-1. The CDE apparently does not interfere with basal transcription from the core promoter (Ref.24). Its function is dependent on a stretch of upstream sequences that is needed for transcriptional activation (UAS). This led to the hypothesis that CDF-1 may function by regulating the activity of upstream activators in a cell cycle-dependent fashion (Ref.24, Ref.52). This conclusion is supported by the observation that the proteins interacting with the cdc25C UAS not only bind constitutively in vivo, but in a heterologous context also activate transcription in a way that is not significantly influenced by the cell cycle (Ref.52).

The major transactivator of the cdc25C UAS is the transcription factor CBF/NF-Y, which binds to three sites 5' to the CDE (Ref.52). A second important transactivator is Sp1 (or other members of the Sp family), which interacts with two sites further upstream (Ref.52). Interestingly, the major activation domains in both Sp1 (Ref.2) and NF-Y (Ref.21, Ref.4) are glutamine-rich, and both factors are therefore likely to contact a similar set of basal transcription factors, TAFs or other components of the preinitiation complex (Ref.10). It cannot, however, be ruled out at present that CDF-1 simply functions by preventing protein contacts through steric hindrance (Ref.16). It is therefore of obvious importance to consider the question as to whether this repression mechanism is restricted to certain class(es) of activation domains (Ref.27), and thus activator-basal complex contacts (Ref.10).

The relevance of this question is stressed by the fact that the promoters of various other cell cycle genes, such as cdc2 and cyclin A, are also regulated through CDE/CHR elements and harbor multiple Sp1 and NF-Y sites in the UASs (Ref.54, Ref.51). In addition, cell cycle genes repressed by the transcription factor E2F also show a conspicuous preference

for Sp1 and NF-Y sites 5' to the E2F sites (Ref.51). It is therefore likely that the molecular basis for E2F- and CDE-mediated negative regulation is very similar and that the repression of glutamine-rich activators like Sp1 and NF-Y is a common mechanism of cell cycle-regulated transcription.

WO96/06943 describes the DNA element present in control sequences of the human cdc25C gene and other cell cycle regulated genes, and proposes the use of the element as a promoter for the expression of genes in a cell cycle regulated manner, e.g. for use in gene therapy of cells undergoing uncontrolled proliferation. WO96/06943 also proposes that one or more proteins bind to this sequence and that such proteins may be isolated by biochemical purification procedures followed by affinity chromatography using an immobilised, multimeric repressor binding site.

Disclosure of the invention.

We have investigated the activity of CDF-1. Although in normal cells CDF-1 activity does not appear to be regulated by protein-protein interaction, we have now found that it is a target for viral oncoproteins. In particular, both SV40 large T antigen and the human papilloma virus E7 gene product appear to disrupt protein interaction in the CDE-CHR *in vivo*.

This finding provides the basis for novel assays useful in the development of modulator compounds suitable for modifying the interaction of such viral proteins with CDF-1 or its cognate binding sequence. Such compounds are useful in the study of mechanisms of viral transformation of cells, and for the development of novel therapeutics.

Furthermore, our findings have allowed CDF-1 to be purified to a high degree of purity for the first time. This is because previous attempts to purify this factor have failed due to the presence of the E7 protein in HeLa cells. HeLa cells are a preferred cell line for the isolation of human transcription

factors because they can be grown in spinner cultures on a large scale. However the presence of E7 from HPV18 in HeLa cells suppresses CDF-1 binding to its cognate DNA binding site and thus makes it impossible to monitor purification through multi-step purification procedures used to isolate proteins of this type. Thus purification has only been possible under conditions in which the effect of E7 on CDF-1 is disrupted, e.g. through the presence of a detergent.

Accordingly, the present invention provides an assay for a modulator of viral replication, wherein said replication is facilitated by suppression of promoter repression mediated by CDF-1 due to the action of a viral protein, which assay comprises:

a) bringing said viral protein into contact with one or both of:

(i) a CDF-1 protein; and

(ii) a DNA sequence comprising a CDF-1 binding site,

in the presence of a putative modulator under conditions wherein in the absence of modulator the viral protein is able to induce inhibition of CDF-1 binding to DNA; and

b) measuring the degree of modulation of binding of CDF-1 to DNA.

The assay is particularly useful when said viral protein is a human papilloma virus E7 protein, and in such an assay the E7 protein or E7 induced inhibitor may be brought into contact with a CDF-1 protein. Other viral proteins include viral oncoproteins which are structurally homologous to E7 such as SV40 large T antigen.

Suitable assay formats include an electrophoretic mobility shift assay (EMSA) to determine the binding of CDF-1 to a CDF-1 binding site, measuring the activity of a reporter gene, said gene being operably linked to a promoter which comprises a CDF-1 binding site, or a two hybrid assay.

In a further aspect, the invention provides a purified CDF-1

protein preparation wherein said CDF-1 is enriched at least 10,000 fold over cellular components with which it is naturally associated. Such a preparation is obtainable by:

a) providing a nuclear extract of mammalian cells;
b) precipitating said extract in the presence of 30% ammonium sulphate;
c) dialysing said precipitate at pH7.5 in the presence of 60mM NaCl to resuspend a CDF-1 fraction;
d) passing said fraction through a first DEAE anion exchange column using an NaCl gradient of 50-1000 mM and collecting the 100-130 mM fraction;
e) passing said fraction from (d) through an EDF anion exchange column using an NaCl gradient of 120-1000 mM and collecting the 120-180 mM fraction;
f) centrifuging the fraction of (e) through a 5-20% glycerol gradient and collecting the 5-7% fraction;
g) passing the fraction of (f) through a cation exchange column using an NaCl gradient of 100-1000 mM and collecting the fraction above 800mM;
h) purifying the fraction of (g) using sequence specific DNA affinity purification using a DNA capture probe comprising a CDF-1 binding site; and
i) recovering the CDF-1.

The invention further provides isolated fragments of the CDF-1 which retains the ability to bind to the cdc25 promoter CDF-1 binding site.

Such fragments, as well as the CDF-1 protein, may be used to provide antibodies capable of binding to CDF-1, and these antibodies form a further aspect of the invention.

The provision of purified CDF-1 allows cDNA encoding CDF-1 to be isolated, for example by a method which comprises probing an expression library with an antibody of the invention, and selecting a clone which expresses a polypeptide capable of binding said antibody. In addition to such cDNA, the invention also provides nucleic acids which have at least 70%

homology to said cDNA, including nucleic acid which encodes a polypeptide capable of binding to the cdc25 promoter CDF-1 binding site.

The present invention also provides an expression system encoding the CDF-1 protein or fragments thereof binding to the CDE-CHR nucleotide sequence or fragments hereof binding to E7 of papilloma virus.

In addition, the present invention provides an expression system encoding an antibody or an antibody fragment (e.g. Fab, Fv, scFv) that mimic the binding site of the CDF-1 molecule for the CDE-CHR nucleotide sequence or for the E7 molecule or protein induced by the E7 molecule or mimics the binding site of the E7 molecule or protein induced by the E7 molecule for the CDF-1 molecule.

Such expression systems may be useful in modulating the cycle of the target cell, into which the expression system has been transferred i.e., it may block the cell cycle promoting activity of E7 or may stimulate cell proliferation by inhibiting CDF-1.

Such effects may be useful to inhibit proliferation of cells infected with HPV or to enhance cell proliferation in cell culture systems or in disease caused by lack of cell proliferation, e.g. decubitus and chronic wounds in diabetes.

The present invention also provides an expression system that encodes a transcription factor which is a fusion protein of an activation domain, a binding site for E7 or CDF-1 and a DNA-binding domain, whereby the transcription factor is under the control of a first activation sequence and binds to a second activation sequence controlling the expression of a structural gene.

The binding site for CDF-1 may be E7 or an antibody or an antibody fragment specific for CDF-1, e.g., for the binding

site of CDF-1 for E7 or molecule induced by E7.

The binding site for E7 may be CDF-1 or an antibody or an antibody fragment specific for E7, e.g., for the binding site of E7 or molecule induced by for CDF-1.

Binding of CDF-1 or E7 to that binding site may inhibit the transcription factor and by that mechanism may also inhibit the expression system.

Thus, the expression system is controlled by the presence in the cell of free CDF-1 or of free E7.

Description of the Drawings.

Figure 1 shows luciferase activity (RLU) of cdc25C reporter gene constructs co-transfected into G0 or S/G2 NIH3T3 cells with empty vector (pMo), E7 wild type (E7) or E7 mutants (see examples).

Figure 2A shows CDF-1 purified from NIH3T3 nuclear extracts. Figure 2B shows CDF-1 purified from HeLa nuclear extracts.

Figure 3 shows the structure of the CDE-CHR oligonucleotides used for sequence-specific DNA affinity purification of CDF-1.

Figure 4 shows the SV40 enhancer, but not the CMV enhancer, is partially repressed in G_0 cells through the CDE-CHR in the context of the cdc25C core promoter.

Figure 5 shows repression through the CDE-CHR is functional with heterologous core promoters.

Figure 6 shows cell cycle regulation of luciferase constructs containing pentameric binding sites for different transcription factors linked to the CDE/CHR and the cdc25C core promoter (-20 to + 121).

Figure 7 shows cell cycle regulation of a luciferase construct containing 5 copies of a Gal4 binding site linked to the CDE/CHR and the cdc25C core promoter (-20 to + 121), as indicated at the top of the Figure, after co-transfection of various transactivation domains fused to a Gal4 fragment harboring the DNA-binding domain (Ref.12, Ref.13).

Figure 8 shows activation potential of transactivation domains fused to a Gal4 fragment harboring the DNA-binding domain after co-transfection with a luciferase construct containing 5 copies of a Gal4 binding site linked to the CDE/CHR and the cdc25C core promoter (-20 to + 121; as in Fig. 7) in growing cells.

Figure 9 shows repression through the CDE-CHR is dependent on its orientation.

Figure 10 shows structure-function analysis of the cdc25C CHR.

Figure 11 shows sequences of cyclin A promoter mutants.

Figure 12 shows effects on cell cycle kinetics of specific nucleotide changes in the cyclin A CDE.

Detailed description of the invention.

Definitions.

A putative modulator compound may be one which enhances, stabilizes or inhibits the various interactions described herein, and thus reference to "modulation" includes both enhancement, stabilization or inhibition of said interactions. Preferably, modulator compounds will be those which inhibit viral replication.

Reference to "viral protein" is to be understood to encompass cellular proteins which are virally induced or modified by the action of proteins encoded by the genome of a virus, such that

said induction or modification is responsible for the effect on CDF-1 activity.

The term "viral replication" is used generally to encompass the various stages of a viral life cycle. The interaction between the viral protein and CDF-1 or its cognate binding sequence may thus facilitate maintenance of any part of this cycle, be it uncoating, replication of genetic material, packaging or other effects mediated on a cell by viral infection, including the effects of viral oncoprotein expression caused by integration of viral genes into the host cell genome, even in the absence of the entire viral genome.

The action of large T antigen and E7 protein on CDF-1 indicates that this protein may be a target for viral proteins in general, particularly oncoproteins including those with pocket protein binding function. Thus cells infected with other viruses, for example adenoviruses, herpes viruses, retroviruses, hepatitis viruses with RNA (e.g. HAV or HCV) or DNA genomes (e.g. HBV), and the like may be examined for the presence of viral proteins which bind to CDF-1 or its cognate binding site.

The provision of purified CDF-1 allows such tests to be carried out in a number of ways known per se in the art and within the remit of those of ordinary skill. The effect of viral proteins on CDF-1 may be examined using methods analogous to those described in the accompanying examples showing the effect of E7 on EMSA assays. Where the presence of viral proteins are determined to have an effect on the activity of CDF-1 these may be investigated further in a variety of ways.

In a preferred aspect the viral protein is the HPV E7 protein and also cellular proteins which are virally induced or modified by the action of proteins encoded by the genome of HPV. The term "E7 protein" includes an E7 protein from any type of HPV, particularly HPV1, 5, 6, 11, 16 or 18. The

sequence of the E7 protein from these or other strains may be obtained by reference to the literature (e.g. Seedorf *et al* Virology, 1985, 145;181-185 for HPV 16) or to computer databases such as Genbank. The term also covers variants and fragments which retain the ability of the protein to bind to CDF-1 as determined by either EMSA or the ability of the fragment or variant to abrogate repression of the cdc25C promoter in G_0 cells. Such fragments or variants may be made using routine genetic engineering techniques and tested in a routine manner based on the exemplification of the accompanying examples.

E7 proteins may also exist in the form of fusion proteins, for example to provide E7 with a detectable tag (such as glutathione S-transferase) in pull down assay formats, a DNA binding or activation domain for the provision of two-hybrid assay formats, or a epitopic or chelateable tag such as a HSV-tag epitope or a polyhistidine tag (such as found in pET vectors available from Novagen, Madison WI, USA - see also US Patent 4,952,496) to allow for purification of the protein from an *in vitro* expression system.

E7 proteins may also be labelled with a detectable label such as biotin or a radiolabel including ^{35}S, ^{32}P or ^{121}I.

The term "CDF-1 binding site" or "cognate binding site" and the like is the CDE-CHR module described above with reference to the prior art. The sequence of this is highly conserved between the promoters of genes repressed by CDF-1, primarily *cdc25C* and *cdc2*. The wild type sequence of *cdc25C* CDF-1 binding site is 5' **GGCGGAAGGTTTGAA** 3' (SEQ ID NO:1) with the most essential residues indicated in bold. Generally, at least some 5' and/or 3' flanking sequence will also be included - various examples of such sequences are included in the examples. Non-wild type sequences which retain activity and are recognised by the wild type CDF-1 protein may also be used.

Assays

Assays according to the invention may be performed in a variety of formats either *in vitro* or *in vivo*. Many suitable formats are know to those of skill in the art for the detection of an interaction between two proteins. A number of suitable formats are described herein although the invention is not limited to these formats alone. The precise format of the assay of the invention may be varied by those of skill in the art using routine skill and knowledge.

For example, the interaction between a viral protein and CDF-1 may be studied by labelling one with a detectable label and bringing it into contact with the other which has been immobilized on a solid support. Suitable detectable labels include ^{35}S-methionine which may be incorporated into one of other protein when produced recombinantly. The recombinantly produced protein may also be expressed as a fusion protein containing an epitope which can be labelled with an antibody.

The protein which is immobilized on a solid support may be immobilized using an antibody against that protein bound to a solid support or via other technologies which are known *per se*. A preferred *in vitro* interaction may utilize a fusion protein including glutathione-S-transferase (GST). This may be immobilized on glutathione agarose beads. In an *in vitro* assay format of the type described above the putative modulator compound can be assayed by determining its ability to modulate the amount of labelled viral protein or CDF-1 protein which binds to the immobilized GST-viral protein or GST-CDF-1, as the case may be. This may be determined by fractionating the glutathione-agarose beads by SDS-polyacrylamide gel electrophoresis. Alternatively, the beads may be rinsed to remove unbound protein and the amount of protein which has bound can be determined by counting the amount of label present in, for example, a suitable scintillation counter.

Alternatively an antibody attached to a solid support and directed against one of the viral protein or CDF-1 may be used in place of GST to attach the molecule to the solid support. Antibodies against the viral protein and CDF-1 may be obtained in a variety of ways known as such in the art, and as discussed herein.

In an alternative mode, one of the viral protein and CDF-1 may be labelled with a fluorescent donor moiety and the other labelled with an acceptor which is capable of reducing the emission from the donor. This allows an assay according to the invention to be conducted by fluorescence resonance energy transfer (FRET). In this mode, the fluorescence signal of the donor will be altered when the viral protein and CDF-1 interact. The presence to a candidate modulator compound which modulates the interaction will increase the amount of unaltered fluorescence signal of the donor.

FRET is a technique known per se in the art and thus the precise donor and acceptor molecules and the means by which they are linked to a viral protein and CDF-1 may be accomplished by reference to the literature.

Suitable fluorescent donor moieties are those capable of transferring fluorogenic energy to another fluorogenic molecule or part of a compound and include, but are not limited to, coumarins and related dyes such as fluoresceins, rhodols and rhodamines, resorufins, cyanine dyes, bimanes, acridines, isoindoles, dansyl dyes, aminophthalic hydrazines such as luminol and isoluminol derivatives, aminophthalimides, aminonaphthalimides, aminobenzofurans, aminoquinolines, dicyanohydroquinones, and europium and terbium complexes and related compounds.

Suitable acceptors include, but are not limited to, coumarins and related fluorophores, xanthenes such as fluoresceins, rhodols and rhodamines, resorufins, cyanines, difluoroboradiazaindacenes, and phthalocyanines.

A preferred donor is fluorescein and preferred acceptors include rhodamine and carbocyanine. The isothiocyanate derivatives of these fluorescein and rhodamine, available from Aldrich Chemical Company Ltd, Gillingham, Dorset, UK, may be used to label the viral protein and CDF-1. For attachment of carbocyanine, see for example Guo et al, J. Biol. Chem., **270**; 27562-8, 1995.

Assays of the invention may also be performed *in vivo*. Such an assay may be performed in any suitable host cell, e.g a bacterial, yeast, insect or mammalian host cell. Yeast and mammalian host cells are particularly suitable.

To perform such an assay *in vivo*, constructs capable of expressing the viral protein and CDF-1 and a reporter gene construct may be introduced into the cells. This may be accomplished by any suitable technique, for example calcium phosphate precipitation or electroporation. The three constructs may be expressed transiently or as stable episomes, or integrated into the genome of the host cell.

In vivo assays may also take the form of two-hybrid assays wherein the viral protein and CDF-1 are expressed as fusion proteins, one being a fusion protein comprising a DNA binding domain (DBD), such as the yeast GAL4 binding domain, and the other being a fusion protein comprising an activation domain, such as that from GAL4 or VP16. In such a case the host cell (which again may be bacterial, yeast, insect or mammalian, particularly yeast or mammalian) will carry a reporter gene construct with a promoter comprising a DNA binding elements compatible with the DBD. The reporter gene may be a reporter gene as disclosed herein. The promoters for the genes may be those discussed herein. Two-hybrid assays may also be conducted *in vitro*, for example in cell-free expression systems (such as a rabbit reticulocyte system) in which the viral protein and CDF-1 protein are expressed in the presence of a reporter gene and the modulator.

Constructs capable of expression of the viral protein and CDF-1 and the reporter gene, may be introduced into the cell and expressed transiently or stably.

Alternatively, assays of the invention may be conducted by utilizing the ability of the viral protein to mediate the activation of a promoter comprising a CDF-1 binding site during the G_0 phase of the cell cycle. The promoter may be operably linked to a reporter gene such as chloramphenicol acetyl transferase, luciferase, green fluorescent protein and β-galactosidase. In this aspect of the invention, modulator and a reporter gene may be introduced into cells which are in G_0, wherein said cells are transformed by a virus which expresses the viral protein or a construct which expresses said viral protein. In the latter case this may be via the native promoter for the protein or from a heterologous promoter.

Where the cells of the assay express CDF-1 then the assay may utilise the natural expression of this protein. This will generally be the case for mammalian cells, such as murine or other rodent, more particularly primate, especially human cells. However such cells may also be modified by targeted deletion of the CDF-1 gene such that said gene is then supplied to the cell in the form of a construct comprising a promoter operably linked to nucleic acid encoding CDF-1. The promoter may be the native CDF-1 promoter or may be a heterologous promoter, such as a CMV promoter.

An alternative assay format relies on the ability of the viral protein to induce disruption of binding of CDF-1 to its cognate DNA binding site, thus abolishing or reducing the electrophoretic mobility shift observed when CDF-1 and DNA comprising said binding site are brought into contact under conditions in which CDF-1 is normally capable of binding to this site (i.e. physiological conditions). In this assay format one or other, or both, of the CDF-1 and its cognate binding site may be labelled with a detectable label and the

electrophoretic mobility of the DNA may be observed. Where the viral protein is present and its ability to bind CDF-1 is not modulated, it abolishes the mobility shift observed, or reduced the signal observed of the shifted band relative to a control. The utility of this assay format is specifically exemplified in the accompanying examples which shows that a detergent is capable of disrupting the interaction between E7 or E7 induced inhibitor and CDF-1.

There is also evidence that CDF-1 interacts with transcription factors, particularly Sp1 and NF-Y which have cognate binding sites in the cdc25C promoter. The provision of CDF-1 allows the present invention to provide assays for modulators of CDF-1 supression of gene expression by providing CDF-1 and a transcription factor, particluarly a Sp1 or NF-Y, or both, optionally in the presence of nucleic acid comprising binding sites for CDF-1 and/or the transcription factors, in the presence of a putative modulator, and determining the effect of the modulator on the interaction between CDF-1 and the transcription factor. Such an interaction may be assessed directly, e.g in a two hybrid assay format as discussed elsewhere herein, or by determining the effect of CDF-1 activity on its cognate binding site wherein said site is operably linked to a reporter gene.

Assays of the invention may also be provided to determine the phosphorylation status of CDF-1, or to determine modulators which affect the phosphorylation status. In general terms, the assays may take any covenient format which is suitable to determine the phosphorylation status of CDF-1, e.g. by using labeled phosphate, or by EMSA, since the binding of CDF-1 to the CDE-CHR site is dependent upon phosphorylation.

Thus such assays include an assay for a modulator which comprises providing

(i) CDF-1 in a phosphorylated form; and

(ii) a putative modulator,

under reaction conditions in which CDF-1 is, in the absence of

the modulator, subjected to a dephosphorylation reaction; and determining the effect of the modulator on the dephosphorylation reaction. The reaction conditions may comprise supplying a phosphatase.

Similarly the invention provides an assay for a modulator of cell cycle progression which comprises providing

(i) CDF-1 in a dephosphorylated form; and

(ii) a putative modulator,

under reaction conditions in which CDF-1 is, in the absence of the modulator, subjected to a phosphorylation reaction; and determining the effect of the modulator on the phosphorylation reaction. The reaction conditions may comprise supplying a kinase.

The assays may be performed using a preparation of CDF-1 which is either dephosphorylated (e.g. from S/G_2 cells) or phosphorylated (e.g. from G_0 cells).

Suitable kinases include the G_1 kinases such as cyclinE-cdk2 or cyclinD-cdk4. Suitable phosphatases include cdc25A and B phosphatases, and bacterial alkaline phosphatase. The phosphorylation status of CDF-1 may be determined by the amout of binding to the CDE-CHR binding site, or for example, by using a labelled source of phosphate and measuring incorporation of the label into the CDF-1.

Modulator compounds suitable for use in assays of the present invention may be any compound available in the art with putative or potential utility. Such compounds include small peptides which are based upon the domains of the viral proteins and/or CDF-1 which interact with each other or intermediary proteins, as such peptides may act as antagonists of the interaction between the two proteins.

The concentration of modulator compounds used in assays will vary depending upon the nature of the compound and the assay format, although generally concentrations will be in the

nanomolar to millimolar range, e.g. from 10 nM to 10 mM, such as from 0.1 μM to 100μM.

Assays of the invention have a number of utilities, including the development of assay systems for the study of cell cycle control as well as for the development of potential therapeutic compounds. Therapeutic compounds which may be developed include those for use in conditions associated with viral transformation of cells. For example, HPV, particularly HPV 16 or HPV18 transformation of cervical cells is associated with the development of cervical cancer and thus the assay of the present invention which utilise an E7 protein of HPV 16 may be used for the development of compounds useful in the treatment of cervical cancer. Where such assays utilise human cells, cell lines derived from cervical cancer cells may be particularly suitable.

More generally, cell lines, particularly human cell lines, which either are transformed by a particular virus which expresses a viral protein which interacts with CDF-1 (or untransformed cells of the same type) may be used for the development of assays for potential therapeutic compounds which are suitable for the treatment of diseases associated with proliferation of that particular cell type.

CDF-1

The invention also provides methods for producing CDF-1 in highly purified form, resulting from the surprising finding that the presence of E7 in HeLa cell preparations causes CDF-1 to be in a form which means that the purification of this can not be monitored unless the interaction between CDF-1 and viral protein is disrupted. The disruption may be achieved by the presence of a detergent in the purification steps. This has allowed for the first time purified protein to be obtained. We have determined that the CDF-1 preparation prepared in accordance with the invention provides a polypeptide having a molecular weight by SDS-PAGE of about 80

kDa.

Thus the invention provides a purified preparation of CDF-1 which in which the CDF-1 is enriched about 10,000 fold in comparison to cellular components with which it is naturally associated with, for example in a human cell such as a HeLa cell.

CDF-1 may be characterised by being obtainable by the multi-step process (a) to (j) set out above. Those of skill in the art will appreciate that this multi-step process may be varied within normal experimental boundaries. Furthermore, although the process provides a first way to enable the provision of CDF-1 and thus its recombinant production, CDF-1 manufactured by other processes, particularly recombinant production, will have the same amino acid structure and is thus also provided by the invention.

In step (a), preferred mammalian cells are human cells, although other cells,e.g. primate or murine cells, or tissues, e.g. brain tissue, are also suitable.

In step (b), we have found that CDF-1 activity appears in a 30% ammonium sulphate precipitation. However, this does not exclude the use of other values, partiuclarly in the range of 20-40%, e.g. 25-35% ammoniumn sulphate.

In step (c), we have found dialysis against 60mM NaCl to be most suitable, although the salt concentration range may be varied, for example between 50 and 100 mM. In addition, other salts may be used. While a pH of 7.5 has been used by us, some variation to this will be acceptable, for example to maintain a pH in the range 7.0 to 8.0.

In step (d), a preferred DEAE (diethylaminoethyl) anion exchange column is a DEAE-agarose (Sepharose) column, although DEAE-dextran (Sephadex), and DEAE-Cellulose (Sephacel), columns may be used. A particularly preferred column is a

DEAE Sepharose Fast Flow (FF) column (Pharmacia), or an equivalent column with a particle size of about 45-165 μm, exclusion limit of $4x10^6$.

In step (e), EDF ion exchangers are cross-linked polymethacrylate resin bonded with a functional group of diethylaminoethyl via linear polymer chains. Preferred ion exchangers of this type have a particle size of 20-40μm, and a capacity of 100 mg/ml BSA.

In step (f), glycerol gradient centrifugation was preformed as described in the accompanying examples. However, the speed and time of the centrifugation to achieve fractionation of the CDF-1 fraction and the number of fractions collected may be varied by those of skill in the art. Generally, it is desirable that the conditions produce a gradient from which a number of fractions, n, are collected such that the CDF-1 activity appears in 10-15% of the fractions (e.g. 3/24 in our case).

In step (g), a preferred cation exchange column is a strong cation exchange column with sulfonic functional groups. Such columns may have an average particle size of 10 μm and a capacity of about 75 mg/ml IgG. The column we have used is a MonoS (Pharmacia) column with functional methly sulphonate groups.

In step (h), the capture probe comprising the CDF-1 binding site we have used in illustrated in Figure 3. This probe comprises the cdc25C promoter binding site, although other sites comprising the CDE-CHR modules described herein may also be used, for example the sites from the cdc2 or cyclin A promoters. The purification of step (h) includes the presence of a detergent in step (h) when the cell extract is from HeLa cells or other virally infected cells which contain a viral protein which interacts with CDF-1. Suitable detergents include NP-40 and/or sodium deoxycholate. The concentration of the detergent may be around from 0.05 to 5%, such as from

0.5 to 2%.

The provision of the purified CDF-1 allows the production of antibodies against this protein. The term "antibody" should be construed as covering any binding substance having a binding domain with the required specificity. Thus the invention covers antibody fragments, derivatives, functional equivalents and homologues of antibodies, including synthetic molecules and molecules whose shape mimics that of an antibody enabling it to bind an antigen or epitope.

Example antibody fragments, capable of binding an antigen or other binding partner are the Fab fragment consisting of the VL, VH, Cl and CH1 domains; the Fd fragment consisting of the VH and CH1 domains; the Fv fragment consisting of the VL and VH domains of a single arm of an antibody; the dAb fragment which consists of a VH domain; isolated CDR regions and F(ab')2 fragments, a bivalent fragment including two Fab fragments linked by a disulphide bridge at the hinge region. Single chain Fv fragments are also included.

Humanized antibodies in which CDRs from a non-human source are grafted onto human framework regions, typically with the alteration of some of the framework amino acid residues, to provide antibodies which are less immunogenic than the parent non-human antibodies, are also included within the present invention

A hybridoma producing a monoclonal antibody according to the present invention may be subject to genetic mutation or other changes. It will further be understood by those skilled in the art that a monoclonal antibody can be subjected to the techniques of recombinant DNA technology to produce other antibodies or chimeric molecules which retain the specificity of the original antibody. Such techniques may involve introducing DNA encoding the immunoglobulin variable region, or the complementarity determining regions (CDRs), of an antibody to the constant regions, or constant regions plus

framework regions, of a different immunoglobulin. See, for instance, EP-A-184187, GB-A-2188638 or EP-A-0239400. Cloning and expression of chimeric antibodies are described in EP-A-0120694 and EP-A-0125023.

Hybridomas capable of producing antibody with desired binding characteristics are within the scope of the present invention, as are host cells, eukaryotic or prokaryotic, containing nucleic acid encoding antibodies (including antibody fragments) and capable of their expression. The invention also provides methods of production of the antibodies including growing a cell capable of producing the antibody under conditions in which the antibody is produced, and preferably secreted.

The reactivities of antibodies on a sample may be determined by any appropriate means. Tagging with individual reporter molecules is one possibility. The reporter molecules may directly or indirectly generate detectable, and preferably measurable, signals. The linkage of reporter molecules may be directly or indirectly, covalently, e.g. via a peptide bond or non-covalently. Linkage via a peptide bond may be as a result of recombinant expression of a gene fusion encoding antibody and reporter molecule.

One favoured mode is by covalent linkage of each antibody with an individual fluorochrome, phosphor or laser dye with spectrally isolated absorption or emission characteristics. Suitable fluorochromes include fluorescein, rhodamine, phycoerythrin and Texas Red. Suitable chromogenic dyes include diaminobenzidine.

Other reporters include macromolecular colloidal particles or particulate material such as latex beads that are coloured, magnetic or paramagnetic, and biologically or chemically active agents that can directly or indirectly cause detectable signals to be visually observed, electronically detected or otherwise recorded. These molecules may be enzymes which

catalyse reactions that develop or change colours or cause changes in electrical properties, for example. They may be molecularly excitable, such that electronic transitions between energy states result in characteristic spectral absorptions or emissions. They may include chemical entities used in conjunction with biosensors. Biotin/avidin or biotin/streptavidin and alkaline phosphatase detection systems may be employed.

The mode of determining binding is not a feature of the present invention and those skilled in the art are able to choose a suitable mode according to their preference and general knowledge.

The provision of the purified protein also allows the gene encoding the protein to be cloned. For example, the gene may be cloned by preparing peptide fragments of purified CDF-1, sequencing said fragments and designing probes based on the genetic code to regions of low redundancy. Reference may be made to standard text books such as Sambrook *et al*, (Cold Spring Harbor, 1989) and the like for these type of techniques. The probes may be used to probe a cDNA or genomic DNA library, in the case of the cDNA library made from cells known to express CDF-1. Clones which hybridize to the probes may be isolated and sequenced. Where an incomplete reading frame is obtained, RACE-PCR methods may be used to provide longer clones from mRNA if desired using standard protocols and kits available in the art. RACE-PCR methods may also be used directly on mRNA to clone the gene where protein sequencing allows probes of sufficient length and sufficiently low redundancy to be made.

Partial protein sequence data also allows EST databases to be searched and ESTs which correspond (i.e. encode) peptides derived from CDF-1 may be obtained and used, if they contain less than a full open reading frame, a probes to clone CDF-1.

Alternatively the provision of antibodies allows expression

screening methods using for example lambda gt11 libraries, wherein clones which express a protein product detectable by the antibodies of the invention may be selected, sequenced and used if need be to obtain further clones. In particular, the finding that *in vitro* phosphorylation by cyclin E-cdk2 can bring deactivated CDF-1 (in a dephosphorylated form in S/G_2 cells) binding to its cognate DNA binding site, made south west screening of cDNA expression libraries (e.g. lambda gt11) possible. Thus, after transferring bacteriophage plaques from a library to a cellulose membrane, *in vitro* phosphorylation of expressed protein by cyclinE-cdk2 is performed. This allows the phosphorylation of CDF-1 expressed in the bacteria, and as a consequence CDF-1-encoding clones are able to bind a CDE-CHR probe. Differential screening of replica plates, from treated and untreated lifts may be performed, allowing rapid identification of positive clones.

A further alternative method provided by the present invention for the cloning of CDF-1 is a two hybrid approach in which a viral protein is labelled with, for example, a DNA binding domain, and is used a bait against an expression library in which clones are linked to an activation domain sequence, such that clones encoding a portion of CDF-1 capable of interacting with the viral protein will activate a suitable reporter construct. E7 may be used as the viral protein for this method.

The knowledge of the size of the CDF-1 protein and its DNA binding properties allows positive checks to be made during the cloning procedure so that clones whose complete open reading frame does not encode a protein of about 80 kDa (as judged by SDS-PAGE and allowing for any glycosylation) and which do not interact with a CDE-CHR may be disregarded.

The provision of native CDF-1 encoding nucleic acid allows allelic and species variants of this gene to be made, as well as synthetic variants. Thus the present invention provides nucleic acid variants and fragments of CDF-1 which, unless the

context is explicitly to the contrary, are included by the term "CDF-1" in relation to the assays of the invention described above, and the various vector and expression means described below.

Nucleic acid includes DNA (including both genomic and cDNA) and RNA, and also synthetic nucleic acids, such as those with modified backbone structures intended to improve stability of the nucleic acid in a cell. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones, addition of acridine or polylysine chains at the 3' and/or 5' ends of the molecule. For the purposes of the present invention, it is to be understood that the polynucleotides described herein may be modified by any method available in the art. Such modifications may be carried out in order to enhance the *in vivo* activity or lifespan of polynucleotides of the invention. Where nucleic acid according to the invention includes RNA, reference to the sequences shown in the accompanying listings should be construed as reference to the RNA equivalent, with U substituted for T.

Nucleic acid of the invention may be single or double stranded polynucleotides. Single stranded nucleic acids of the invention include anti-sense nucleic acids.

Generally, nucleic acid according to the present invention is provided as an isolate, in isolated and/or purified form, or free or substantially free of material with which it is naturally associated, such as free or substantially free of nucleic acid flanking the gene in the human genome, except possibly one or more regulatory sequence(s) for expression. Nucleic acid may be wholly or partially synthetic and may include genomic DNA, cDNA or RNA.

Nucleic acids of the invention include nucleic acids having at least 70% sequence identity to native human CDF-1 obtainable by the methods described above. Preferably the degree of

sequence identity in either case is at least 80%, such as at least 90%, 95%, 98% or 99%.

The invention also provides nucleic acids which are fragments of the nucleic acids described in the preceding paragraph. Particular nucleic acids fragments which are preferred include:

(a) nucleic acids which comprise a sequence encoding an active portion of the invention, said active portion having the ability to bind to the cdc25C CDF-1 binding site; and

(b) nucleic acids which consist essentially of from 15 to 50, for example from 15 to 35, 18 to 35, 15 to 24, 18 to 30, 18 to 21 or 21 to 24 nucleotides of a sequence having at least 70% homology to the nucleic acids described in the preceding paragraph.

The term "consist essentially of" refers to nucleic acids which do not include any additional 5' or 3' nucleic acid sequences. In a further aspect of the invention, nucleic acids of the invention which consist essentially of from 15 to 30 nucleotides as defined above may however be linked at the 3' but preferably 5' end to short (e.g from 4 to 15, such as from 4 to 10 nucleotides) additional sequences to which they are not naturally linked. Such additional sequences are preferably linkers which comprise a restriction enzyme recognition site to facilitate cloning when the nucleic acid of the invention is used for example as a PCR primer.

Nucleic acids of the invention, particularly short (less than 50) sequences useful as probes and primers may carry a revealing label. Suitable labels include radioisotopes such as ^{32}P or ^{35}S, fluorescent labels, enzyme labels, or other protein labels such as biotin. Such labels may be added to polynucleotides or primers of the invention and may be detected using by techniques known per se.

Also included within the scope of the invention are antisense

sequences based on the nucleic acid sequences described herein, preferably in the form of oligonucleotides, particularly stabilized oligonucleotides, or ribozymes. Antisense oligonucleotides may be designed to hybridise to the complementary sequence of nucleic acid, pre-mRNA or mature mRNA, interfering with the production of polypeptide encoded by a given DNA sequence (e.g. either native CDF-1 polypeptide or a mutant form thereof), so that its expression is reduce or prevented altogether. The construction of antisense sequences and their use is described in Peyman and Ulman, Chemical Reviews, 90:543-584, (1990), Crooke, Ann. Rev. Pharmacol. Toxicol., 32:329-376, (1992), and Zamecnik and Stephenson, P.N.A.S, 75:280-284, (1974). The construction of ribozymes and their use in described in for instance Gibson and Shillitoe, Molecular Biotechnology 7(2): 125-137, (1997).

Nucleic acid sequences encoding all or part of the *CDF-1* gene and/or its regulatory elements can be readily prepared by the skilled person using the information and references contained herein and techniques known in the art (for example, see Sambrook, Fritsch and Maniatis, "Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989, and Ausubel et al, Short Protocols in Molecular Biology, John Wiley and Sons, 1992). These techniques include (i) the use of the polymerase chain reaction (PCR) to amplify samples of such nucleic acid, e.g. from genomic sources, (ii) chemical synthesis, or (iii) preparing cDNA sequences. Modifications to the wild type sequences described herein can be made, e.g. using site directed mutagenesis, to lead to the expression of modified polypeptides or to take account of codon preference in the host cells used to express the nucleic acid.

In general, short sequences for use as primers will be produced by synthetic means, involving a step wise manufacture of the desired nucleic acid sequence one nucleotide at a time. Techniques for accomplishing this using automated techniques are readily available in the art.

Longer polynucleotides (including those from subjects expressing inactive portions) will generally be produced using recombinant means, for example using a PCR (polymerase chain reaction) cloning techniques. This will involve making a pair of primers (e.g. of about 15-50 nucleotides) based on the sequence information provided herein to a region of the mRNA or genomic sequence encoding the mRNA which it is desired to clone, bringing the primers into contact with mRNA or cDNA obtained from a murine or human cell (e.g. a brain cell, particularly a fetal brain cell), performing a polymerase chain reaction under conditions which bring about amplification of the desired region, isolating the amplified fragment (e.g. by purifying the reaction mixture on an agarose gel) and recovering the amplified DNA. The primers may be designed to contain suitable restriction enzyme recognition sites so that the amplified DNA can be cloned into a suitable cloning vector.

Such techniques may be used to obtain all or part of the nucleic acid of the invention. Genomic clones containing the *CDF-1* gene and its introns and promoter regions may also be obtained in an analogous manner, starting with genomic DNA from a murine or human cell, e.g. a primary cell such as a liver cell, a tissue culture cell or a library such as a phage, cosmid, YAC (yeast artificial chromosome), BAC (bacterial artificial chromosome) or PAC (P1/P2 phage artificial chromosome) library.

Polynucleotides which are not 100% homologous to the sequences of the present invention but fall within the scope of the invention can be obtained in a number of ways.

Other human variants (for example allelic forms) of the *CDF-1* gene described herein may be obtained for example by probing cDNA or genomic DNA libraries made from human tissue.

In addition, other animal, for example fish (such as the Zebra fish), worm (such as *C.elegans*) and particularly mammalian

(e.g. rodent, rabbit, sheep, goat, pig, or primate) homologues of the *CDF-1* gene may be obtained. Such sequences may be obtained by making or obtaining cDNA libraries made from dividing cells or tissues or genomic DNA libraries from other animal species, and probing such libraries with probes comprising all or part of a nucleic acid of the invention under conditions of medium to high stringency: for example for hybridization on a solid support (filter) overnight incubation at 42°C in a solution containing 50% formamide, 5xSSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulphate and 20 μg/ml salmon sperm DNA, followed by washing in 0.03M sodium chloride and 0.03M sodium citrate (i.e. 0.2x SSC) at from about 50°C to about 60°C).

Alternatively, such polynucleotides may be obtained by site directed mutagenesis of the native human CDF-1 gene. This may be useful where for example silent codon changes are required to sequences to optimise codon preferences for a particular host cell in which the polynucleotide sequences are being expressed. Other sequence changes may be desired in order to introduce restriction enzyme recognition sites, or to alter the property or function of the polypeptides encoded by the polynucleotides. Further changes may be desirable to represent particular coding changes which are required to provide, for example, conservative substitutions.

In the context of cloning, it may be necessary for one or more gene fragments to be ligated to generate a full-length coding sequence. Also, where a full-length encoding nucleic acid molecule has not been obtained, a smaller molecule representing part of the full molecule, may be used to obtain full-length clones. Inserts may be prepared from partial cDNA clones and used to screen cDNA libraries. The full-length clones isolated may be subcloned into expression vectors and activity assayed by transfection into suitable host cells, e.g. with a reporter plasmid.

The present invention also extends to nucleic acid comprising transcriptional control sequences for the *CDF-1* gene. Such control sequences will be found 5' to the open reading frame of the *CDF-1* gene and are obtainable by probing a genomic DNA library (such as a phage, cosmid, YAC, BAC or PAC library) of a mammal with a nucleic acid of the invention, selecting a clone which hybridizes under conditions of medium to high stringency, and sequencing the clone 5' to the open reading frame of the gene. Where only a small amount of sequence is present in the 5' region, this sequence may be used to reprobe the library to genome walk further upstream. Analysis of the upstream region will reveal control regions for gene expression including control regions common to many genes (i.e TATA and CAAT boxes) and other control regions, usually located from 1 to 10,000, such as 1 to 1000 or 50 to 500 nucleotides upstream of the start of transcription.

To confirm that such regions are the control regions of the gene, they may be linked to a reported gene (such as β-galactosidase) and tested in any suitable *in vitro* or *in vivo* system. For example the construct of the control region (e.g. comprising 50 to 500 nucleotides upstream of the start of transcription) and the reporter gene may be used to produce a transgenic animal (see below) and the pattern of expression, both spatially and developmentally, may be compared with that of the *CDF-1* gene. Where substantially similar patterns of expression are found, this shows that the construct comprises substantially all of the control region of the wild type gene.

The control region may be mutated to identify specific subregions responsible for transcriptional control. This may be achieved by a number of techniques well known in the art as such, including DNase protection footprint assays, in which the control region is brought into contact with an extract from a cell in which the *CDF-1* gene is actively expressed, and the regions of the control region which bind factors in that extract is determined.

Isolated nucleic acid comprising such control regions obtainable by such a method form a further aspect of the present invention.

The present invention further extends to genomic DNA exon sequences found between the introns encoding a *CDF-1* gene in an animal subject, such as those mentioned above and including humans. Such exon sequences may be obtained in a manner analogous to that described above for the transcriptional control sequences, with the appropriate genome walking being conducted between the intron sequences. The locations of the exons may be determined by comparing genomic and cDNA sequences of the *CDF-1* gene, observing where the sequences line up and diverge, and looking for consensus splice sequences which define intron/exon boundaries.

Exon sequences obtainable by these or analogous methods may be used in the construction of mini-gene sequences which comprise nucleic acid encoding polypeptides of the invention which sequences are interrupted by one or more exon sequences.

Mini-genes may also be constructed using heterologous exons, derived from any eukaryotic source.

Nucleic acid according to the present invention, such as a full-length coding sequence or oligonucleotide probe or primer, may be provided as part of a kit, e.g. in a suitable container such as a vial in which the contents are protected from the external environment. The kit may include instructions for use of the nucleic acid, e.g. in PCR and/or a method for determining the presence of nucleic acid of interest in a test sample. A kit wherein the nucleic acid is intended for use in PCR may include one or more other reagents required for the reaction, such as polymerase, nucleosides, buffer solution etc. The nucleic acid may be labelled. A kit for use in determining the presence or absence of nucleic acid of interest may include one or more articles and/or reagents for performance of the method, such as means for providing the

test sample itself, e.g. a swab for removing cells from the buccal cavity or a syringe for removing a blood sample (such components generally being sterile). In a further aspect, the present invention provides an apparatus for screening nucleic acid, the apparatus comprising storage means including the a nucleic acid or the invention or fragment thereof, the stored sequence being used to compare the sequence of the test nucleic acid to determine the presence of mutations.

Polynucleotides or primers of the invention or fragments thereof labelled or unlabelled may be used by a person skilled in the art in nucleic acid-based tests for detecting the *CDF-1* gene in the human or animal body. In the case of detecting, this may be qualitative and/or quantitative. Detection includes analytical steps such as those which involve sequencing the gene in full or in part.

Such tests for detecting generally comprise bringing a human or animal body sample containing DNA or RNA into contact with a probe comprising a polynucleotide or primer of the invention under hybridizing conditions and detecting any duplex formed between the probe and nucleic acid in the sample. Such detection may be achieved using techniques such as PCR or by immobilizing the probe on a solid support, removing nucleic acid in the sample which is not hybridized to the probe, and then detecting nucleic acid which has hybridized to the probe. Alternatively, the sample nucleic acid may be immobilized on a solid support, and the amount of probe bound to such a support can be detected. Suitable assay methods of this any other formats can be found in for example WO89/03891 and WO90/13667.

In one embodiment, the sample nucleic acid may be in the form of whole chromosomes, for example as a metaphase spread. The nucleic acid probe or primer of the invention may be labelled with a fluorescent label to detect the chromosomal location of a *CDF-1* gene in the spread.

Where a PCR based assay is used in the detection of nucleic

acid according to the invention, it is preferred that both primers are nucleic acids according to the present invention. However, one of the two primers may be directed to sequences which flank sequences of the *CDF-1* gene or are contained within its exons. Those of skill in the art will be able to select specific pairs of PCR primers using routine skill and knowledge in the light of the present disclosure.

A further method of detection according to the invention is in detecting changes to wild-type *CDF-1* genes, including single base changes, using single stranded conformational polymorphism (SSCP) analysis. Nucleic acid sequence from all or part of a *CDF-1* DNA or mRNA in a sample is hybridized to a reference sequence, and the mobility of the hybrid is observed in a gel under conditions where any non-hybridized regions within the duplex give rise to changes in mobility.

Nucleic acids of the invention are thus useful in screening a test sample containing nucleic acid for the presence of alleles, mutants and variants, the probes hybridising with a target sequence from a sample obtained from the individual being tested. The conditions of the hybridisation can be controlled to minimise non-specific binding, and preferably stringent to moderately stringent hybridisation conditions are preferred. The skilled person is readily able to design such probes, label them and devise suitable conditions for the hybridisation reactions, assisted by textbooks such as Sambrook et al (1989) and Ausubel et al (1992).

As well as determining the presence of polymorphisms or mutations in the *CDF-1* sequence, the probes may also be used to determine whether mRNA encoding the *CDF-1* gene is present in a cell or tissue.

Nucleic acid of the invention may be provided in the form of compositions, for example a pharmaceutical composition. Such compositions will include pharmaceutically acceptable carriers and adjuvants. Examples of a suitable carrier include

liposomes. Liposomes carrying nucleic acid of the invention (particularly where such nucleic acid is carried by a vector, see below) may be used in methods of gene delivery in gene therapy. Suitable liposome compositions and delivery systems are described in Gill *et al*, Gene Therapy, Vol.4, No.3, pp.199-209 (1997).

Nucleic acid of the invention can be incorporated into a recombinant replicable vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus in a further embodiment, the invention provides a method of making polynucleotides of the invention by introducing a polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about replication of the vector. The vector may be recovered from the host cell. Suitable host cells are described below in connection with expression vectors.

Preferably, a CDF-1 polynucleotide of the invention in a vector is operably linked to a control sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector.

The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A control sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under condition compatible with the control sequences.

Suitable vectors can be chosen or constructed, containing appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate. Vectors may be plasmids, viral e.g. 'phage phagemid or baculoviral, cosmids, YACs, BACs, or PACs as appropriate. Vectors include gene therapy vectors, for

example vectors based on adenovirus, adeno-associated virus, retrovirus (such as HIV or MLV) or alpha virus vectors.

The vectors may be provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a neomycin resistance gene for a mammalian vector. Vectors may be used *in vitro*, for example for the production of RNA or used to transfect or transform a host cell. The vector may also be adapted to be used *in vivo*, for example in methods of gene therapy. Systems for cloning and expression of a polypeptide in a variety of different host cells are well known. Suitable host cells include bacteria, eukaryotic cells such as mammalian and yeast, and baculovirus systems. Mammalian cell lines available in the art for expression of a heterologous polypeptide include Chinese hamster ovary cells, HeLa cells, baby hamster kidney cells, COS cells and many others.

Promoters and other expression regulation signals may be selected to be compatible with the host cell for which the expression vector is designed. For example, yeast promoters include S. cerevisiae GAL4 and ADH promoters, S. pombe nmt1 and adh promoter. Mammalian promoters include the metallothionein promoter which is induced in response to heavy metals such as cadmium. Viral promoters such as the SV40 large T antigen promoter or adenovirus promoters may also be used. All these promoters are readily available in the art.

The vectors may include other sequences such as promoters or enhancers to drive the expression of the inserted nucleic acid, nucleic acid sequences so that the polypeptide is produced as a fusion and/or nucleic acid encoding secretion signals so that the polypeptide produced in the host cell is secreted from the cell.

Vectors for production of polypeptides of the invention of for use in gene therapy include vectors which carry a mini-gene sequence of the invention.

For further details see, for example, Molecular Cloning: a Laboratory Manual: 2nd edition, Sambrook et al., 1989, Cold Spring Harbor Laboratory Press. Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in Current Protocols in Molecular Biology, Ausubel et al. eds., John Wiley & Sons, 1992.

Vectors may be transformed into a suitable host cell as described above to provide for expression of a polypeptide of the invention. Thus, in a further aspect the invention provides a process for preparing polypeptides according to the invention which comprises cultivating a host cell transformed or transfected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptides, and recovering the expressed polypeptides. Polypeptides may also be expressed in in vitro systems, such as reticulocyte lysate.

A further embodiment of the invention provides host cells transformed or transfected with the vectors for the replication and expression of polynucleotides of the invention. The cells will be chosen to be compatible with the said vector and may for example be bacterial, yeast, insect or mammalian.

Polynucleotides according to the invention may also be inserted into the vectors described above in an antisense orientation in order to provide for the production of antisense RNA or ribozymes.

A still further aspect of the present invention provides a

method which includes introducing the nucleic acid into a host cell. The introduction, which may (particularly for *in vitro* introduction) be generally referred to without limitation as "transformation", may employ any available technique. For eukaryotic cells, suitable techniques may include calcium phosphate transfection, DEAE-Dextran, electroporation, liposome- or polycation-mediated transfection and transduction using retrovirus or other virus, e.g. vaccinia or, for insect cells, baculovirus. For bacterial cells, suitable techniques may include calcium chloride transformation, electroporation and transfection using bacteriophage. As an alternative, direct injection of the nucleic acid could be employed.

The introduction may be followed by causing or allowing expression from the nucleic acid, e.g. by culturing host cells (which may include cells actually transformed although more likely the cells will be descendants of the transformed cells) under conditions for expression of the gene, so that the encoded polypeptide is produced. If the polypeptide is expressed coupled to an appropriate signal leader peptide it may be secreted from the cell into the culture medium. Following production by expression, a polypeptide may be isolated and/or purified from the host cell and/or culture medium, as the case may be, and subsequently used as desired, e.g. in the formulation of a composition which may include one or more additional components, such as a pharmaceutical composition which includes one or more pharmaceutically acceptable excipients, vehicles or carriers (e.g. see below).

A further aspect of the present invention provides a host cell containing nucleic acid as disclosed herein. The nucleic acid of the invention may be integrated into the genome (e.g. chromosome) of the host cell. Integration may be promoted by inclusion of sequences which promote recombination with the genome, in accordance with standard techniques. The nucleic acid may be on an extra-chromosomal vector within the cell.

Nucleic acids of the invention, particularly when in the form

of a recombinant vector, may be used in methods of gene therapy. A construct capable of expressing a nucleic acid of the invention may be introduced into cells of a recipient by any suitable means, such that a polypeptide of the invention, preferably a full length or active portion polypeptide, is expressed in the cells.

The construct may be introduced in the form of naked DNA, which is taken up by some cells of animal subjects, including muscle cells of mammalians. In this aspect of the invention the construct will generally be carried by a pharmaceutically acceptable carrier alone. The construct may also formulated in a liposome particle, as described above.

The following examples illustrate the invention.

EXAMPLE 1: VIRAL ONCOPROTEINS DISRUPT THE FUNCTION OF CDF-1

In this example we show that the S/G_2-specific cdc25C promoter is deregulated in a similar way as E2F-regulated genes, i.e. by the oncoprotein-mediated dissociation of a repressor. We first demonstrated this question by genomic (in vivo) footprinting of human fibroblasts (WI-38 cells) transformed by the SV40 large T oncoprotein (WI-38 VA13 cells).

Normally growing WI-38 and WI-38 VA13 cells were sorted by FACS after DMS treatment and fixation. This procedure yielded almost pure fractions of G_1 and G_2 cells. As previously reported, protection of the CDE in cdc25C promoter cells was seen in the G_1 cells, and this occupation was lost in G_2. In WI-38 VA13 cells, however, the cdc25C CDE was unprotected in both G_1 and G_2, although the interactions with constitutively bound proteins in the upstream region were unaffected by expression of SV40 large T (positions -27/-28, -39 to -44, -58/-59, -70, -91 to -96, -103/-105).

In order to confirm and extent these observations we also analyzed protection of the CDE in the cyclin A promoter. It

has been previously been shown that viral oncoprotein expression leads to transcriptional deregulation of the cyclin A gene. The cyclin A CDE differs from the cdc25C CDE, in that it can interact with both CDF-1 and E2F (see Example 4 below). The principal repressor of the cyclin A promoter is CDF-1, while E2F binding appear to affect mainly the timing of CDE derepression (see Example 5). The genomic footprint showed that in the control cells four G nucleotides (-34, -32, -31, and -30) of the CDE core sequence were protected in G_1, and that this protection was lost in cells in G_2. In contrast, there was no detectable protection of the CDE throughout the cell cycle in WI-38 VA13 cells. Additional constitutive protections of G nucleotides were visible in both cell lines (-52, -53,-64,-77, and -80 relative to the major start site of transcription), as described previously, confirming the validity of the results.

Taken together, these observations clearly suggest that the large T oncoprotein disrupts protein interaction in the CDE-CHR region in vivo.

To confirm this effect of large T protein in a functional assay, an expression plasmid encoding large T (and the corresponding empty vector) was transfected into NIH3T3 cells together with the cdc25C promoter-driven luciferase reporter gene C290 (Ref.24, Ref.54). Cells were synchronized in G_0 and luciferase activities were determined as described (Ref.24). The results clearly showed that cotransfection of the large T vector led to a >10-fold increase in cdc25C promoter activity in G_0 cells. Since cell cycle progression was only marginally affected ($\leq$2-fold; as shown by FACS analysis of cells cotransfected with a EGFP vector and stained for DNA content), these results indicate that large T directly derepresses the cdc25C promoter.

Very similar results were also obtained with an expression vector encoding the human papilloma virus E7 protein cotransfected with C290 (see Fig. 1). E7 is structurally and

functionally similar to SV40 large T antigen and adenovirus E1A. There are three domains highly conserved in these viral oncoproteins, conservered region (CR) 1, CR2 and CR3. In E7, the CR2 domain is required for its association with pRB family members, pRB, p107 and p130. This interaction abrogates the ability of pRB family members to bind and to inactivate transcription factor E2F. The CR2 domain is thus crucial to the oncogenic function of E7. The CR1 domain of E7 is required for cellular transformation, induction of epidermal hyperplasia and tumour formation in transgenic mice. It has also been observed that the disruption of CR1 impairs anchorage-independent growth of rodent fibroblasts which is normally induced by E7. To identify the essential domains of E7 required for the induction of cdc25C , we tested three mutants with amino acid changes in the CR1 (mutant Pro2) and the CR2 (mutants Gly24 and Gly26) domains. The CR1 Pro2 position has been shown to be required for E7 oncogenicity, but not for E7 binding to pRB family members. The Gly24 mutant is unable to interact with pRb and p107, while Gly26 is defective in pRb binding, but able to interact with p107. As can be seen in Fig. 1, the two mutants of CR2 were unable to abrogate repression of the cdc25C promoter in G_0 cells, which suggests that the CR2 domain plays a key role in the deregulation of the cdc25C promoter. Interestingly, the CR1 mutant Pro2 also shows the defect in induction of the cdc25C promoter in G_0 cells. The significance of this discovery is the CDF-1 is the first identified cellular protein whose inactivation requires CR1 and contributes to the cellular transformation of E7 function.

Finally, we analyzed the effect of E7 protein on the DNA-binding properties of CDF-1 in vitro. To this end, we performed EMSA with nuclear extracts from NIH3T3 cells stably transformed by E7 and the corresponding control cells harboring the empty expression vector. NIH3T3 cells were cultured in Dulbecco-modified Eagle medium (DMEM) with 10% fetal calf serum. For synchronization in G_0, cells were maintained in serum free medium for 2 days. HeLa cells were

grown in S-MEM plus 5% newborn calf serum. NIH3T3 cells were transfected by the DEAE dextran technique EMSA was performed as described in Example 4 using a cdc25C-wt probe. As can be seen in Fig. 2A, a clear CDF-1-DNA complex was detectable in the control cells, but not in the E7-transformed cells. However, upon addition of detergent (0.8% Na-deoxycholate, 1.5% NP-40), CDF-1 complex formation could be restored in the presence of E7. These detergents are known to disrupt protein-protein interactions, such as pocket protein-E2F interactions, and thus presumably liberate CDF-1 from a complex with E7 or an E7-induced inhibitor of DNA-binding. Comparable results were also obtained with HeLa cells which are known to express E7 protein (Fig 2B). A CDF-1-DNA complex could only be detected in the presence of detergent, a finding that is also highly relevant to the purification of CDF-1 (see Example 2).

The experimental setup described here (EMSA) can be used for the identification of compounds that disrupt the interaction of CDF-1 with the inhibitory activity in E7-transformed cells and thus have a potential value for the treatment of diseases caused by human papilloma viruses. This approach appears is of particular interest, because CDF-1 activity does not appear to be regulated by protein-protein interactions during normal cell cycle progression (unlike E2F which is regulated by pocket proteins). This is clearly suggested by EMSAs showing that the DNA-binding properties of CDF-1 in normal S/G_2 cells cannot be restored by detergent treatment. Therefore, a specific inhibitor of CDF-1 protein interactions should not affect normal cell physiology.

EXAMPLE 2: PURIFICATION OF CDF-1

This example describes the novel purification route enabling, for the first time, highly purified CDF-1 to be obtained.

2.1 Ammomium sulfate precipitation

Hela cell nuclear extract (250 ml; 5 to 10 mg/ml of protein)

was precipitated by the step-wise addition of solid ammonium sulfate at 4°C to final concentrations of 0, 30, 40, 50, and 60% saturation, respectively (Methods in Enzymology, vol 182, pp. 291). At each step, the mixture was centrifuged at 20,000 rpm for 30 min (Beckmann JA-20 rotor). The precipitated protein was redissolved in 5-20 ml of buffer A (50 mM Tris•Cl pH 7.5, 1 mM EDTA, 10% glycerol, 3mM DTT) plus 100 mM NaCl and 0.1% NP-40, and analyzed by EMSA for the presence of CDF-1. The bulk of CDF-1 activity appeared in the 30% fraction. For further purification, removal of ammonium sulfate was accomplished by dialysis against buffer A plus 60 mM NaCl for at least 4 hours.

2.2. Anion exchange chromatography

The CDF-1 containing fraction obtained after ammonium sulfate precipitation was centrifuged to remove any precipitates and purified by two successive anion-exchange chromatography steps: DEAE-Sepharose fast flow (Pharmacia) and Fractogel EMD DEAE-650(s) (EDF; Merck). DEAE-Sepharose 60 ml was equiliberate with buffer A plus 60 mM NaCl and packed into a Pharmacia XK26 column. The elution was carried out at a flow rate of 1.5 ml/ min using a NaCl gradient of 50-1000 mM. The CDF-1 activity appeared at 100-130 mM NaCl. The CDF-1 fractions from three DEAE purifications were collected, diluted with 50 mM Tris•Cl pH 7.5 (0.33x volume) and applied to an EDF column. Elution was carried out at a flow rate of 1 ml/min using a NaCl gradient of 120-1000 mM plus 0.1% NP-40. The 120 to 180 mM NaCl fractions were directly used for the next stept purification.

2.3. Glycerol gradient purification

Two mililiters of the eluate from the EDF column was applied to 10 ml gradients of 5 to 20% glycerol in buffer B (50 mM Tris•Cl pH 7.5, 1 mM EDTA, 10% glycerol, 3mM DTT) plus 100 mM NaCl. The proteins were separated by centrifugation at 40,000 rpm for 22 hours at 4°C (Beckmann SW-40 rotor). Samples were collected in 24 fractions starting at the top of the gradient and analysed by EMSA. The CDF-1 activity appeared in the top

fractions (4 to 6) only (Mr <100 kD).

2.4. Cation exchange chromatography

A strong cation exchanging column-Mono S (1 ml; Pharmacia) was used for the further purification. Samples from two glycerol gradient purifications (about 30 ml) were loaded onto the column. The elution gradient was from 100 mM to 1 M NaCl (total volume of 20 ml) at a flow rate of 0.5 ml/min. The CDF activity did not elute from the Mono S column at NaCl concentration of <800 mM which indicates that it contains a highly positively charged (basic) domain.

2.5. Sequence-specific DNA affinity purification

Biotin-labeled top-strand oligonucleotide (100 nmol; Fig. 3 SEQ ID NO:2) and the complementary unlabeled oligonucleotide (250 nmol; Fig. 3 SEQ ID NO:3) were mixed in a buffer containing 67 mM Tris•Cl pH7.6, 13 mM $MgCl_2$, 6.7 mM DTT, 1.3 mM spermidine, and 1 mM EDTA in a total volume of 1 ml. This mixture was incubated at 95°C for 5 min, 65°C for 20 min, 37°C for 20 min and room temperature for 10 min. The annealed oligonucleotide duplex was coupled to 400 μl of streptavidin agarose (GIBCO-BRL) equiliberated with TE buffer (10 mM Tris•Cl pH 7.5, 0.1 mM EDTA) by incubation at room temperature for 15 min. The coupled streptavidin agarose was then equilibrated in a Bio-Rad Econo-Column with buffer Z (25 mM Hepes (K^+), pH7.8, 12.5 mM $MgCl_2$, 1 mM DTT, 20% (vol/vol) glycerol and 0.1% Nonidet P-40) containing 0.1 M KCl. The fractions eluted from the Mono S column were mixed with buffer Z, sonicated salmon sperm DNA (0.05 mg/ml), sodium deoxycholate (final concentration 0.8%) and NP-40 (final concentration 1.5%) and loaded to the streptavidin agarose affinity column. The the column was washed with 10 ml of buffer Z containing 0.1 M KCl at a flow rate of 0.25 ml/min. Elution was performed by the successive addition of 1 ml of buffer Z containing 0.2, 0.25, 0.3, 0.4, 0,6 and 1 M KCl, respectively, directly into the column and thorough mixing with a siliconized glass rod. For each step, the streptavidin agarose was allowed to settle for 5 min prior to elution. The

CDF-1 binding activity appeared i the 0.3 and 0.4 M KCl fractions. The same fractions specifically contained a protein of ~80 kDa that could by visualized SDS-PAGE followed by by siver-staining. In agreement with this observation, a ~80 kDa protein was crosslinked by UV-light to the CDE-CHR probe.

EXAMPLE 3: CDF-1-MEDIATED REPRESSION OF CELL CYCLE GENES TARGETS A SPECIFIC SUBSET OF TRANSACTIVATORS

In this example we have used chimeric promoter constructs to demonstrate that the cdc25C UAS, but not the core promoter, is crucial for repression. In addition, we show that only specific transcription factors and activation domains are responsive to CDE/CHR-mediated cell cycle regulation. These observations clearly indicate that CDF-1 interferes with the activation of transcription by a specific subset of transactivators, particularly glutamine-rich activators, pointing to specific interactions of CDF-1 with components of the transcription machinery. In agreement with this conclusion we find that a simple inversion of the CDE-CHR module completely abrogates cell cycle-regulated repression.

Materials and Methods

Cell culture, DNA transfection and luciferase assays

NIH3T3 cells were cultured in Dulbecco-Vogt modified Eagle medium (DMEM) supplemented with 10% fetal calf serum, penicillin and streptomycin. Cells were transfected by the DEAE dextran technique. For synchronization in G_0, cells were maintained in serum free medium for 2 days after transfection. Determination of luciferase activities were performed as published. Mel-ab cells (Ref.46) (provided by Prof. I. Hart, London) were grown in DMEM plus 10% FCS, 200 nM TPA and 10mM cholera toxin. C2C12 cells (Ref.47) (ATCC CRL-1772; obtained from Prof. H.H. Arnold, Braunschweig) were grown in DMEM plus 20% FCS. Synchronization of cells in G_0/G_1 was followed by FACS analysis as described (Ref.53).

Plasmid constructs

The cdc25C promoter-luciferase constructs C290, C75, C33 and C20 as well the CDE mutants R1 and T7 (referred to here as mCDE) and the CHR mutant -6/-3 (referred to here as mCHR) have been described elsewhere (Ref.24, Ref.54, Ref.52). Other constructs were generated by cloning of synthetic oligonucleotides with appropriate terminal overhangs or by PCR strategies as previously described (Ref.54, Ref.52, Ref.9). The following oligonucleotides were used (as 5 copies) for cloning of reporter constructs containing multiple transcription factor binding sites: NF-Y (Ea-Y; MHC class II promoter (Ref.25): 5'-ATTTTTCTGATTGGTTAA (SEQ ID NO:4);NFk-B (mouse k light chain enhancer (Ref.41): 5'-AGAGGGGACTTTCCGAGA (SEQ ID NO:5); NF-1/CTF (high affinity binding site from adenovirus origin of replication (Ref.3): 5'-TTTTGGCTACAAGCCAATA (SEQ ID NO:6); Sp1: 5'-ATGGGGCGGAGA (SEQ ID NO:7)(Ref.2). Gal4: 5'CGGAGTACTGTCCTCCG (SEQ ID NO:8) (Ref.5). The oligonucleotides were sythesized with BamHI and BglII termini and cloned into the BamHI-digested cdc25C promoter construct C20 (Ref.24). The Gal4 expression vectors are based on pGal(1-147) (Ref.43) (GAYA plasmids; see Table 1) or pCG (Ref.5) (all other constructs). Chimeric promoters were generated by fusing the following enhancers or UASs to the cdc25C constructs C20 (Ref.24): human troponin C promoter from -98 to +23 (Ref.34), human myf-4/myogenin promoter from -210 to +54 (Ref.32), human tyrosinase enhancer (-2.0 to -1.8 kb; EcoRI-NcoI fragment) (Ref.37), mouse TRP-1 promoter from -332 to -23 (Ref.23), SV40 promoter/enhancer from -281 to -45 (genomic region 273 to 36) (Ref.48) and CMV enhancer (292 to 695 region form pcDNA3, Invitrogen). In the case of SV40 and CMV enhancers, fusions were also made with the cdc25C constructs C33, C51 and C20/+30 (Ref.24). For inversion of the CDE/CHR in the cdc25C promoter the sequence 5'-GGGCTGGCGGAAGGTTTGAAT (SEQ ID NO:9) was changed to 5'-GttcaaaccttgccT (SEQ ID NO:10). Other constructs were cloned by using PCR-generated promoter fragments as indicated in the Results and the Figures. All PCR-amplified fragments were verified by DNA sequencing using the dideoxynucleotide

chain-termination method using Sequenase 2.0 (USB) or Tth polymerase (Pharmacia).

Table 1

Type of Activation Domain	Name of Plasmid	Description
Gln-rich	Gal4-N/Oct pSCTEV GAL4 Sp1:Q2 GAYA-6	Oct-2 (aa 3-15) Sp1 (aa 340-485) NFY-A (aa1-132 D26-53)
Ser-, Thr-, Gln-rich	GAYA-5 GAYA-11	NFY-A (aa 1-233 D26-53) NFY-A (aa1-233)
Ser, Thr-rich	pSCTEVGAL4ITF2:ST	ITF2 (aa2-452)
Acidic	pCG GAL4 VP16 pGAL4 Myc	VP16 (aa 413-390) c-Myc (aa 1-262)
Prol-rich	pCG GAL4 CTF	CTF (aa 399-499)
Empty Vector	pCG (1-94) pGal4 (1-147)	

The cdc25C CDE-CHR module fails to repress heterologous enhancers.

We first addressed the question whether CDE/CHR-mediated repression might be dependent on a specific UAS. For this purpose, the cdc25C UAS was exchanged with various heterologous enhancer sequences. Since it could not be excluded that the spacing between activators and repressor might be important, the heterologous sequences were fused with various cdc25C promoter fragments starting at -51, -33 or -20, and extending to +121. In addition, we used a fragment lacking most of the down-stream sequence, extending to +30. All fragments were used in the wild-type form and, as controls, with mutated CDEs. The cdc25C promoter fragments were fused with both the cytomegalovirus (CMV) enhancer and the simian virus 40 (SV40) early enhancer/promoter regions and tested for cell cycle-regulated repression in transient luciferase assays in NIH3T3 cells.

The results are shown in Figure 4, which has legends as

follows: SV40-TATA and CMV-TATA: natural SV40 and CMV regulatory sequences containing the viral early promoter/enhancer region, TATA-box and transcription start site. Chimeric SV40 and CMV fusion constructs consist of the SV40 early promoter/enhancer region or the CMV enhancer linked to a minimal cdc25C promoter fragment harboring either a wild-type (wt) or mutated CDE (R1) and the core promoter. C20: cdc25C sequence from -20 to +121; C20/+30: -20 to +30; C33: -33 to +121; C51: -51 to +121. The C75 (wild-type) and C75R1 (mutated CDE) cdc25C promoter constructs (-75 to +121) were included for comparison. All constructs were tested in quiescent (G_0) and growing cells. Data are represented as the ratio of luciferase activity in growing versus G_0 cells, normalised to 1 for SV40-TATA (C75 and SV40 fusion constructs) or CMV-TATA (chimeric CMV constructs). Values are given as averages ± standard deviation calculated from 2-4 independent sets of data.

The chimeric CMV constructs did not show any significant cell cycle regulation, in contrast to the chimeric SV40 constructs. In the latter case, cell cycle regulation was, however, only partially restored and approximately 3-fold below the induction value seen with the natural cdc25C promoter construct C290.

We also used the enhancer sequences from a number of other promoters in similar experiments. The human tyrosinase enhancer (Ref.37) and tyrosinase-related protein-1 (TRP-1) UAS (Ref.23) was fused to a cdc25C promoter fragment (-20 to +121). The resulting chimeric constructs were tested in the melanocytic cell line Mel-ab (Ref.46), which can be synchronized in G_0 by serum deprivation. Analogous constructs were generated with the human myf-4/myogenin (Ref.32) and troponin C (Ref.34) UASs and tested in the myoblast cell line C2C12 (Ref.47). These cells were arrested in G_1 by exposure to 5% horse serum instead of 20% fetal calf serum (Ref.32). With none of the constructs were we able to observe any cell cycle regulation. These observation clearly suggest that efficient

cell cycle-regulated repression is observed only with specific activating sequences, such as the cdc25C UAS.

The CDE/CHR module represses the cdc25C UAS in the context of heterologous core promoters.

We next asked the question whether the core promoter might be of similar importance as the UAS. We use the term "core promoter" for the transcription initiation region of the cdc25C gene. This region, which shows basal promoter activity, extends from approximately +1 to +50. The cdc25C UAS plus the CDE/CHR module (-121 to +2) were fused to the core promoters of the human terminal deoxynucleotidyl transferase (TdT) gene (Ref.39) or the SV40 early region (Ref.48). The cdc25C, TdT and SV40 fragments alone, as well as the C290 cdc25C promoter construct (-290 to +121) and the SV40-TATA construct (see legend to Fig. 4) were included for comparison. Black bars: cdc25C sequences; white bars: TdT sequences; hatched bars: SV40 sequences. Data are represented as the ratio of luciferase activity in growing versus G_0 cells. Values are given as averages ±standard deviation calculated from 3 independent experiments. As can be seen in Fig. 5, these chimeric promoters showed a similar cell cycle regulation as the wild-type promoter C290, indicating that CDE/CHR-mediated repression is not core promoter-specific. In the subsequent experiments, we therefore directed our attention to the observed activator specificity.

The cdc25C CDE-CHR module represses the activation by specific transcription factors.

The observations described above suggest that CDE/CHR-mediated repression might work only in the context of specific activators. To address this question more directly we sought to investigate the function of the CDE/CHR module in the context of promoters which interact with only one transcription factor (family). To this end, we constructed promoters where 5 binding sites for NF-Y, NFk-B, NF-I or Sp1 fused to the cdc25C minimal promoter construct C20 (-20 to +121). The constructs were tested in both growing and

quiescent NIH3T3 cells (Fig. 6). The experiment was performed with reporter constructs containing either a wild-type or a mutated CDE to be able to determine the contribution of the CDE/CHR module to cell cycle regulation, expressed as the ratio of the values obtained with wild-type to mutant CDE reporter constructs (right-most column). Data are represented as the ratio of luciferase activity in growing versus G_0 cells. Values are given as averages ±standard deviation calculated from 2 sets of data. All four mutant constructs lacked any significant cell cycle regulation, as did the wild-type ones containing NFk-B or NF-I binding sites. In contrast, CDE/CHR-dependent cell cycle regulated transcription was seen with NF-Y sites (3-fold), and to a lesser extent, with Sp1 sites (2.2-fold). Since NF-Y and Sp1 are the transactivators of the cdc25C promoter (Ref.52), these observations may explain the observed requirement for a specific UAS. The fact that Sp1 is subject to cell cycle regulation via the CDE/CHR module may also offer an explanation for the observation that the SV40 promoter, which contains 6 Sp1 sites (Ref.48), was the only one which could replace the cdc25C UAS without totally obliterating cell cycle-regulated repression (Fig. 4).

The cdc25C CDE-CHR represses specific transactivation domains. In order to analyze the activator specificity of CDE/CHR-mediated repression in more detail we investigated the repressibility of specific transactivation domains fused to a fragment of Gal4 harboring the DNA-binding domain (Ref.5, Ref.13). Plasmid vectors expressing fusion proteins of Gal4 and various activation domains (see Table 1) were co-transfected with a luciferase reporter containing 5 copies of a Gal4 binding site linked to the cdc25C minimal promoter C20 (-20 to +121). The experiment was performed in both growing and G_0 cells with reporter constructs containing either a wild-type or a mutated CDE to be able to determine the contribution of the CDE/CHR module to cell cycle regulation (indicated as ++, +, ± or -). Data are represented as the ratio of the values obtained with wild-type to mutant CDE reporter constructs (as in the right-most column in Fig. 6).

Values are given as averages ± standard deviation calculated from 5-9 independent experiments. The vector control contained only the Gal4 domain without any fused heterologous sequences. Luciferase activities were determined in both growing and quiescent NIH3T3 cells with both wild-type and mutant CDE reporter plasmids (Fig. 7). CDE/CHR-mediated repression was seen with three different activation domains, the glutamine-rich domains of Oct-2 (Oct-N) (Ref.5), Sp1 (Sp1-Q1) (Ref.2) and different fragments of the glutamine-serine-threonine-rich activation domain of NF-Y (GAYA-5, -6 and -11) (Ref.21, Ref.42). The strongest cell cycle regulation was generally seen with GAYA-11, which contains the complete transactivation domain of the A-subunit of human NF-Y (corresponding to the B-subunit of rat CBF) (Ref.4, Ref.38). In contrast, no significant cell cycle-regulated repression was detected with other transactivation domains, i.e., ITF2 (Ref.36), VP-16 (Ref.31) (Ref.5), Myc (Ref.7) and CTF (Ref.5, Ref.26). These domains are not glutamine-rich, and belong to the classes of serine-threonine-rich, acidic or proline-rich transcription factors (Ref.27). These data are in agreement with the results obtained in Fig. 6 and confirm the conclusion that CDE/CHR-mediated repression is specific for a subset of transactivation domains. Significantly, the best repression was seen with NF-Y, which is the major transactivator of the cdc25C gene, and also plays important roles in many other cell cycle genes. The fact that Sp1 (Fig. 6) as well as the Q1 domain of Sp1 (Fig. 7) are less efficiently repressed than NF-Y may, however, not reflect the physiological situation and could rather be due to the artificial experimental setup, but this question is of minor importance with respect to the conclusions drawn from this study. The same applies to the fact that even the NF-Y-based constructs gave rise to a considerably lower cell cycle regulation than the natural cdc25C UAS. In addition, both Sp1 (Ref.2) and NF-Y (Ref.4) contain multiple activation domains, some of undefined nature, which may all contribute to cell cycle regulation in the context of the natural cdc25C promoter.

In general, no correlation was seen between the level of activated transcription and the degree of cell cycle regulation. This is shown in Figure 8 which indicates RLU values are given as averages ± standard deviation calculated from the same experiments as in Fig. 7. The vector control contained only the Gal4 domain without any fused heterologous sequences. Thus, for example, the activation domain showing the best regulation, GAYA-11, gave rise to expression values that were either lower, similar or higher than those seen with the unregulated domains from VP-16, Myc or CTF. Our findings therefore suggest that repressibility (or the lack of it) is an intrinsic property of the activation domains.

Repression by the cdc25C CDE-CHR module is orientation-dependent.

The specificity of repression reported above suggests that the CDE/CHR-interacting factor establishes specific contacts with the transcription machinery. If so, the orientation of the CDE/CHR module should be important for its function. We therefore analyzed a series of constructs where the CDE/CHR module was inverted. C290 and C75-based cdc25C promoter constructs were assayed in both growing and G_0 cells and the ratio of both values was determined. invCDE/CHR: constructs contain inverted CDE/CHR modules as indicated at the top of Figure 9; invCDE: inversion of the CDE only; mCDE and mCDE/CHR: mutations of the CDE or both the CDE and CHR (controls). The SV40-TATA construct (see legend to Fig. 4) was included for comparison. Values are given as averages ± standard deviation calculated from 3 independent experiments. This inversion led to a complete abrogation of cell cycle-regulated repression (without affecting transcription levels in growing cells). The CDE/CHR inversion thus had a similar effect as the mutation of the CDE, its inversion (and thus disruption of the repressor module) or the mutation of both the CDE and CHR (Fig. 9). We have not formally shown that the inverted CDE/CHR still binds CDF-1, but previously published data strongly suggest that the nucleotides flanking

the inverted sequence do not play a role in CDF-1 function (Ref.54). The finding that inversion of the CDE/CHR abrogates cell cycle regulation is therefore in line with the conclusion that CDE/CHR-mediated repression appears to involve stereo-specific interactions with other transcription factors.

The data obtained in the present study suggest that CDE/CHR-mediated repression involves specific protein-protein interactions: (i) the cdc25C UAS could not be replaced by heterologous UASs or enhancers without partially or even completely impairing cell cycle-regulated repression; (ii) the glutamine-rich transactivation domains of NF-Y, Sp1 and Oct-2 were repressed in a cell cycle-dependent manner through the CDE/CHR module while several other activation domains did not show a comparable effect; (iii) inversion of the CDE/CHR sequence in the cdc25C promoter totally abrogated repression.

EXAMPLE 4: CDF-1, A NOVEL E2F-UNRELATED FACTOR, INTERACTS WITH CELL CYCLE-REGULATED REPRESSOR ELEMENTS IN MULTIPLE PROMOTERS.

In this example we have identified the cdc25C CDE-CHR binding activity, the CDE-CHR binding factor-1 (CDF-1). We provide compelling evidence that CDF-1 behaves *in vitro* as expected from the results obtained by *in vivo* footprinting and in functional assays. In addition, we show that CDF-1 also interacts with the CDE-CHR repressor elements in the cyclin A and cdc2 promoters and provide an initial biochemical characterization of this novel activity.

Materials and Methods

Cell culture, DNA transfection and luciferase assays.
NIH3T3 cells were cultured in Dulbecco-modified Eagle medium (DMEM) with 10% fetal calf serum. For synchronization in G_0, cells were maintained in serum free medium for 2 days. HeLa cells were grown in S-MEM plus 5% newborn calf serum. NIH3T3 cells were transfected by the DEAE dextran technique and luciferase activities were determined as described (Ref.24).

Sequence analysis and luciferase constructs.

The cdc25C and B-myb promoter-driven luciferase constructs have been described elsewhere (Ref.24, Ref.54, Ref.22). Mutations were introduced by PCR strategies as previously described (Ref.24, Ref.9).

Protein binding assays.

Electrophoretic mobility shift assays (EMSAs) were performed as described previously (Ref.22). Briefly, nuclear extract (4μg) was incubated in 12μl of a buffer containing 50 mM Tris/HCl (pH8.0), 10% v/v glycerol, 0.2 mM EDTA, 1 mM DTT, 0.8% sodium deoxycholate, and 1μg poly(dA/dT) for 10 min. NP-40 was added to a final concentration of 1.5% and incubation was continued for another 20 min. ^{32}P-labelled probe (0.2 pmol) was added and the reaction mixture was incubated for another 20 min. All reactions were performed on ice. Probes were labelled by filling-in 5' overhanging ends. The following double-stranded probes were used for EMSAs:

cdc25C-wt: 5'-ACTGGGCT**GGCGGAA**GG*TTTGAA*TGGTCAA (SEQ ID NO:11)(CDE bold; CHR italic). T1, T4, T7 (also referred to as cdc25C-mCDE), A8 and C9 are mutated at positions -19, -16, -13, -12 and -11, respectively, as described (Ref.54).

cdc25C -10/-7: 5'-ACTGGGCTGGCGGA*cttg*TTGAATGGTCAA (SEQ ID NO:12)

cdc25C-6/-3 (also referred to as cdc25C-mCHR):

5'-ACTGGGCTGGCGGAAGGT*ggtc*ATGGTCAA (SEQ ID NO:13)

cdc25C-1/+2: 5'-ACTGGGCTGGCGGAAGGTTTGAA*ggt*TCAA (SEQ ID NO:14)

cdc25C-2: 5'-ACTGGGCTGGCGGAAGGTTTGA*c*TGGTCAA (SEQ ID NO:15)

The sequences of all other oligonucleotides, including B-myb, have been described elsewhere (Ref.22). The random oligonucleotide contains an irrelevant sequence (Ref.52). DP-1 antibodies were kindly provided by N. La Thangue (Glasgow). All other antibodies were purchased from Santa-Cruz (SC-251X, SC-632X, SC-879X, SC-512X, SC-999X, SC-830X). *In vitro* dimetyl sulphate (DMS) methylation protection footprinting was performed as described (Ref.22).

Genomic footprinting of stable transfectants.

For the generation of stable cell lines, the wild-type cdc25C luciferase construct C290 and the CHR mutant C290mCHR5/6 (TTTGAA mutated to TagGAA) were inserted into the pAGLu vector (Ref.18) which contains a matrix attachment region (MAR/SAR) and introduced into NIH3T3 cells by electroporation. Stably transfected clones were isolated under G418 selection and analyzed for luciferase expression in quiescent and growing cells. Clones with the expected expression pattern were expanded and analyzed by genomic footprinting (Ref.30) as described (Ref.24) with the exception that the first primer (P1) was specific for the luciferase gene (5'-GTAACACAAAGGAATTCAAGC) (SEQ ID NO:16).

Glycerol gradient centrifugation.

HeLa nuclear extract was dialyzed against a buffer containing 50 mM Tris-HCl, 100 mM NaCl, 1 mM DTT and proteinase inhibitors. Two-hundred microliters of the dialyzed extract were applied to 2 ml gradients of 5% - 20% glycerol and centrifuged at 45,000 rpm for 18 hours. Samples were collected in 20 fractions starting at the top of the gradient. Pocket proteins were identified by immunoblotting using a mixture of antibodies against pRb, p107 and p130 (Santa Cruz; SC-50-G, SC-250X; SC-317X).

Results.

Delineation of the cdc25C CHR.

One important requirement for the identification of the CDF-CHR binding protein(s) is the establishment of correlations between the ability of such proteins to interact with mutated repressor elements and the function of such mutated elements in cell cycle regulation. We have previously defined the consensus sequence of the CDE as G/C G C/T G G/C (GGCGG in cdc25C) (Ref.54). For the CHR, however, such information is not yet available. In order to delineate the borders of the CHR and to identify critical nucleotide positions we introduced a number of point, triple and quadruple mutations into the CHR of the cdc25C promoter and

analyzed the function of these mutant constructs by measuring their repression in NIH3T3 cells synchronized in G_0. The mutations are summarised in Figure 10, together with the results of transient luciferase assays are expressed as the ratio of RLUs observed with growing cells relative to the activity in quiescent cells. The results shown in the Figure summarize the data of 4 independent experiments using at least two independent preparations of plasmid DNA. Values represent averages; standard deviations are indicated by error bars. An SV40 reporter plasmid was included in each experiment to standardize the factor of induction (the SV40 reporter typically gave a 1.5-fold higher value in growing compared to quiescent cells). The data in Fig. 10 clearly show that the CHR extends from -7 to -2, and that all nucleotide positions in this region are crucial. In contrast, the nucleotide positions between the CDE and the CHR (-11 to -8; AAGG) and the nucleotides downstream from the CHR (≥-1; TGG...) can be altered without detectable effects on repressor function. The cdc25C CHR can thus be defined as the sequence TTTGAA.

In vivo CDE occupation is dependent on an intact CHR.

Previous data have clearly shown that CDE and CHR in different promoters function in a synergistic way, since mutations in either element destroy repression in G_0 (Ref.54). This could mean the interacting factor(s) bind(s) cooperatively to both elements. Obviously, this information is of great importance for the biochemical identification of such factor(s). We therefore sought to clarify this question by genomic DMS footprinting of a stably transfected NIH3T3 cell line carrying a cdc25C promoter construct with an inactivating mutation in the CHR (cdc25C-mCHR5/6: TTTGAA changed to TagGAA). The expected protection pattern was observed in a control line stably expressing a wild-type cdc25C promoter construct. In contrast, the cell line harboring the cdc25C promoter with the CHR mutation did not show any protection in the region of the CDE and the mutated CHR, while occupation of two constitutive upstream binding sites for NF-Y (Ref.24) was unchanged in the mutant promoter. We therefore conclude that CDE occupation is

dependent on an intact CHR, indicating cooperative binding *in vivo*. This conclusion is supported by our further observation that the insertion of either 5 bp or 10 bp between the CDE and the CHR in the cdc25C promoter abrogates repression.

In vitro footprinting shows that CDF-1 contacts the CDE in the major groove and the CHR in the minor groove.

In order to obtain additional evidence that CDF-1 is the activity interacting with the repressor elements *in vivo*, we analyzed the interaction of CDF-1 with DNA by DMS methylation protection footprinting *in vitro*. We have previously shown that *in vivo* the CDE is contacted in the major groove, while the CHR is occupied in the minor groove (Ref.54). A very similar result was obtained by *in vitro* footprinting of the upper strand, with specific protection of the four G-residues in the CDE indicating major groove contacts, and of the two A-residues in the CHR indicating minor groove contacts. The mode of interaction between CDF-1 and the CDE-CHR *in vitro* is thus fully compatible with the observations made *in vivo*.

CDF-1 interaction with the CDE/CHR is cell cycle-regulated.

We next sought to establish that the activity identified above shows the expected pattern of cell cycle-regulated DNA-binding. For this investigation we had to switch to NIH3T3 cells, since the cell cycle in HeLa cells is deregulated. NIH3T3 cells were serum-deprived for 48 hrs, restimulated with 10% FCS and nuclear extracts were analyzed by EMSA using the cdc25C CDE-CHR probe in the absence of detergents. CDF-1 activity (identified by its comigration with HeLa CDF-1) in quiescent cells, and at 6 and 12 hrs post-stimulation was observed. At later time points between 16 and 24 hrs, however, CDF-1 binding activity was greatly diminished. This pattern was inversely correlated with the expression of CDC2, whose transcription is controlled by a CDE-CHR repressor module. Thus, the kinetics of CDF-1 binding to DNA are in perfect agreement with its presumptive function as a cell cycle-regulated repressor.

CDF-1 interaction with multiple promoters containing CDE-CHR modules.

Our previous studies have shown that functional CDE-CHR modules are present in different promoters, including cdc25C, cdc2 and cyclin A (Ref.54). In addition, a similar configuration of binding sites is found in the B-myb promoter where an E2F site with a sequence similar to the cdc25C CDE, is located immediately upstream of a CHR-like element (Ref.22, Ref.1). It was therefore of obvious interest to investigate whether the CDF-1 activity identified above would interact with the repressor sites in these promoters. EMSA was carried out with nuclear extract and the cdc25C CDE-CHR as the probe in the presence of various competitors, including the cdc25C CDE-CHR (self-competition), the cyclin A CDE-CHR, the cdc2 CDE-CHR and the B-myb E2FBS-CHR. Each competitor was used as the wild-type sequence and with inactivating mutations in the CDE, E2FBS or CHR (see Materials and Methods for details). The ratio of probe to competitor was in each case 1:20. A random irrelevant oligonucleotide competitor was used as a control. Both CDE-CHR containing promoters, i.e., cdc2 and cyclin A, bind the CDF-1 activity with a similar efficiency as the cdc25C promoter. In all three cases binding was dependent on a cooperative binding to both the CDE and CHR, since mutation in either site impaired competition with the cdc25C probe. At an identical ratio of probe to competitor (1:20), competition by the B-myb promoter E2FBS-CHR module was insignificant, although we also found that some competition could be seen at higher competitor concentrations. The fact that the CDF-1 activity shows a strong and selective interaction with all three CDE-CHR containing promoters provides additional evidence for the relevance of the activity identified in the present study.

CDF-1 does not contain known E2F family members.

In view of the similarity of the CDE with an E2FBS we sought to investigate whether the CDF-CHR activity identified above might contain known E2F or DP family members. For this purpose, EMSA was performed in the presence of antibodies

directed against DP-1, DP-2, E2F-1, E2F-3, E2F-4 and E2F-5 proteins. We have found that all of these antibodies have been shown to either induce supershifts or extinguish binding in different settings (see also Ref.22). However, the results clearly showed that none of the antibodies used affected complex formation, indicating that CDF-1 does not contain any of the known E2F or DP family members. This finding also implies that the cdc25C CDE-CHR sequence does not interact with E2F complexes present in HeLa nuclear extracts. In agreement with this conclusion we could not detect binding of recombinant GST-E2F and GST-DP proteins in any combination of the 8 known family members to the cdc25C CDE-CHR sequence, whereas strong binding to the B-myb E2FBS was observed under the same conditions.

CDF-1 does not co-fractionate with E2F and pocket proteins.

We finally asked the question as to whether CDF-1 might associate with, and thus be regulated by, pocket proteins of the pRb family. For this purpose we separated the nuclear protein (complexes) from HeLa cells under native conditions by glycerol gradient centrifugation. Fractions were analyzed by EMSA for both CDF-1 and E2F binding activities (in the presence of detergent which disrupts higher order complexes) and by immunoblotting for the presence of pocket proteins. E2F binding activity was found in fractions 2-10, while pocket proteins (pRb, p107 and p130) were present in fractions 4-10, but not in fraction 2. Thus, fraction 2 contains "free" E2F, while fractions 4-10 contain the higher order E2F-pocket protein complexes. Strikingly, CDF-1 was detectable to fraction 2 and was not found in fractions 4-10 which contained the pocket proteins. This finding suggests that these proteins may not form complexes with the bulk of CDF-1.

Summary.

The major goal of the present study was to identify the cdc25C CDE-CHR binding activity. Using specific EMSA conditions we were able to identify an activity (CDF-1) that fulfills the criteria expected of a CDE-CHR binding repressor. First, CDF-1

interacts in a cooperative fashion with the CDE and the CHR in the cdc25C promoter which is in agreement with the CHR-dependent occupation of the CDE seen *in vivo*. Second, CDF-1 interacts with G residues in the CDE (major groove) and with A residues in the CHR (minor groove). This protection pattern is identical to the one found by *in vivo* footprinting (Ref.54). Third, the binding of CDF-1 to sequences containing mutated CDE or CHR motifs correlates precisely with the function of such mutated elements in cell cycle regulated repression. Fourth, the DNA-binding activity of CDF-1 is clearly cell cycle-regulated, and is inversely correlated to the expression of the CDC2 gene, whose transcription is controlled by a CDE-CHR repressor module. Fifth, CDF-1 binds with similar efficiency to all known CDE-CHR regulated promoters, i.e., cdc25C, cdc2 and cyclin A, but only weakly to the B-myb E2F site. In each case, this binding was dependent on the presence of intact versions of the CDE and CHR. Taken together these findings are compelling evidence that CDF-1 represents the CDE-CHR interacting repressor.

Our results also suggest that CDF-1 does not form complexes with pocket proteins, unlike E2F. Our observations therefore suggest that the bulk of CDF-1 activity is regulated by other mechanisms, perhaps phosphorylation or the interaction with other factor(s) blocking its DNA-binding domain (see Example 6).

EXAMPLE 5: A NEW MODEL OF CELL CYCLE-REGULATED TRANSCRIPTION: REPRESSION OF THE CYCLIN A PROMOTER BY CDF-1 AND ANTI-REPRESSION BY E2F.

The CDE/CHR in the cyclin A promoter is recognized by both CDF-1 and E2F complexes (Ref.35, Ref.54, Ref.49). It has also been shown that ectopic E2F can induce the cyclin A promoter (Ref.6, Ref.35). However, since the cyclin A promoter is upregulated later than B-myb (which contains a bona fide E2F-BS), but earlier than cdc25C (which interacts only with CDF-1), it is unclear what the actual role of these factors in

the regulation of the cyclin A promoter is. In the present example, we addressed this question mainly by a structure-function approach. By introducing point mutations into the CDE of the cyclin A promoter we were able to generate promoter variants that show a dramatically decreased or increased E2F binding relative to the interaction with CDF-1. These mutants show an unchanged repression in G_0, indicating that E2F binding is dispensable. On the other hand, the loss of CDF-1 binding due to a mutation in the CHR has beeen shown to result in derepression in G_0 (Ref.54). Therefore, CDF-1 is likely to be the principal repressor of the cyclin A gene. The modulation of E2F binding did lead, however, to altered cell cycle kinetics with higher levels of E2F binding causing earlier up-regulation of transcription. Our observations suggest that E2F contributes to the correct timing of cyclin A transcription by acting as an anti-repressor.

Interaction of the cyclin A CDE with E2F.

We first sought to assess the affinity of E2F for the cyclin A CDE by comparing the binding of recombinant E2F-1/DP-1 (rE2F) complex with CDEs or E2F binding sites (E2F-BSs) from other promoters. As expected, a strong interaction was seen with the B-myb E2F-BS in an electrophoretic mobility shift assay (EMSA), which is in agreement with previous data (Ref.53). In contrast, no interaction with rE2F was seen with the cdc25C CDE which interacts specifically with CDF-1. Both, the cyclin A and the cdc2 CDE showed a clear binding to rE2F, but this interaction was considerably weaker as compared to the B-myb E2F-BS. An interaction of both promoter elements with E2F in nuclear extracts has previously been shown (Ref.20, Ref.35, Ref.40, Ref.53, Ref.49), but from the data obtained in the present study we can conclude that the cyclin A CDE does not represent a high affinity E2F-BS. Since the cyclin A CDE-CHR interacts with CDF-1 as efficiently as with the "classical" cdc25C motif (see Example 4), these findings provided a first hint that CDF-1 might play a more important role in repression than E2F complexes.

Abrogation of E2F binding does not affect repression of the cyclin A promoter.

We next analyzed whether an interaction of the cyclin A promoter CDE with E2F complexes is required for the repression of transcription in G_0 cells. To this end, we analyzed a set of CDE mutants for interaction with E2F and CDF-1 in order to screen for promoter variants (Figure 11) with altered E2F binding properties relative to their interaction with CDF-1. Three such variants could be identified: while no E2F binding was detectable with C2A9 and C2G4A9 in the presence of an unchanged interaction with CDF-1, mutant G4 showed a strongly increased affinity for E2F. When these promoter constructs were analyzed in a functional assay in vivo, wild-type-like repression in G_0 was observed with all three constructs (Table 2). These data strongly suggest that the extent or loss of E2F-binding to the cyclin A promoter does not influence its repression in G_0 cells, establishing CDF-1 as the principal repressor of the cyclin A gene.

Table 2: Activities of cyclin A wild-type and mutant constructs in quiescent (G_0) and stimulated (24 hrs) NIH3T3 cells. Values represent the average of 2 experiments.

	G_0 (RLUs/103)	24 hrs (RLUs/103)	Induction (fold)
Cyclin A	0.4 ± 0.01	24.7 ± 2.3	56.3
G4	0.6 ± 0.20	36.3 ± 2.9	55.8
C2A9	0.6 ± 0.02	24.6 ± 0.8	43.0
C2G4A9	0.5 ± 0.07	23.3 ± 1.1	46.6

The cyclin A CDE interacts only weakly with E2F-4, the major family member in G_0 cells.

The observations made above raise the question why the E2F complexes present in G_0/earlyG_1 cells, which contain mainly E2F-4, apparently have no effect on the cyclin A promoter. We therefore analyzed the interaction of various E2F family members with the cyclin A CDE using HeLa cell nuclear extracts

in EMSAs. The individual E2F proteins (E2F-1, -2, -3, -4 and -5, and DP-1 and DP-2) were identified by member-specific antibodies. HeLa cells contained three discernible DNA-binding E2F complexes, E2F-1/DP-1, E2F-3/DP-1 and E2F-4/DP-1, which can, for instance, be visualized with the B-myb promoter E2F-BS (Ref.22). The cyclin A CDE, however, showed a detectable interaction only with E2F-1 and E2F-3 complexes.

We next asked the question why the cyclin A CDE differs in its ability to interact with individual E2F family members from that of the B-myb E2F-BS. This analysis showed that a single nucleotide determines the efficiency of interaction with E2F-4/DP-1: exchanging the first C in the cyclin A CDE to a G (as in B-myb) rendered the cyclin A element able to interact with E2F-4 as efficiently as with E2F-1 or E2F-3.

<u>E2F binding shifts the kinetics of cell cycle regulation.</u>
In order to investigate a potential role for E2F in cyclin A promoter regulation at later stages of the cell cycle, we analyzed the activity of the mutant cyclin A promoter constructs displaying altered affinities for E2F during cell cycle progression of serum-stimulated NIH3T3 cells. NIH3T3 cells were transiently transfected with the indicated constructs, synchronized in G_0 by serum deprivation and stimulated for the different times. The data represent the average of 3 different experiments. Data were normalized to 100 at 24 hrs. for each construct in order to facilitate a comparison of the half-maximal expression values. The kinetics in Fig. 12 demonstrate that all three promoter mutants differed in their patterns of cell cycle regulation from that seen with the wild-type cyclin A promoter. Thus, mutants C2A9 and C2G4A9 were induced later in the cell cycle and in this respect resembled the cdc25C promoter, while mutant G4 was up-regulated earlier than the wild-type promoter, thus mimicking B-myb regulation. These findings strongly suggest that a loss of E2F binding leads to a delay of the onset of transcription, while an increased affinity for E2F results in premature transcriptional up-regulation. The

increased affinity of the G4 mutant for E2F-4 complexes is in perfect agreement with this conclusion, since free E2F-4 is generated in mid-G_1, i.e. prior to E2F-1 and -3 (Ref.33).

These results are in line with the hypothesis that free E2F-1 and -3 complexes, which are found in late G_1, compete with CDF-1 binding to the cyclin A promoter, thus leading to derepression around S-phase entry. The interaction of E2F with the cyclin A CDE might also contribute to promoter activity through transactivation, since the G4 construct showed a slightly increased activity in serum-stimulated cells as compared to the wild-type promoter (Table 2).

Taken together, our observations suggest that the kinetics of cyclin A regulation have three components: (i) the repression by CDF-1 in G_0 and early-mid G_1, (ii) the competition with free E2F in late G_1/S leading to anti-repression and thus inducing the onset of transcription, and (iii) the subsequent inactivation of CDF-1 in S/G_2 (Ref.54) leading to maximal levels of gene expression.

That CDF-1 rather than E2F is the repressor of the cyclin A promoter is suggested by a number of observations: (i) The affinity of E2F for the cyclin A CDE is low compared to promoters harboring E2F-BSs, such as B-myb. (ii) Repression of the cyclin A promoter is critically dependent on the CHR (Ref.52), and changing the sequence of the cyclin A CHR to that of the B-myb promoter abolishes repression. In agreement with this result, the cyclin A CHR is identical to that of the cdc25C gene, which interacts with CDF-1 (see Example 4), but not with E2F. (iii) Mutations of the cyclin A CDE, which lead to a dramatic decrease (or even abolition as in C2A9) of E2F binding without noticeably altering the interaction with CDF-1, do not affect repression of the cyclin A promoter in G_0 cells to any detectable extent. (iv) The cyclin A promoter interacts with E2F-1 and E2F-3 complexes in HeLa and NIH3T3 nuclear extracts, but not with E2F-4 under the identical conditions.

While it seems clear that E2F plays no significant role in repression of the cyclin A promoter in G_0/G_1 cells, several lines of evidence suggest that E2F complexes have a function later in the cell cycle: (i) A clear interaction of the cyclin A CDE could be detected with E2F-1 and E2F-3, which are upregulated in late G_1. At this stage of the cell cycle complex formation of E2F with the CDE is comparable to the binding of CDF-1. (ii) Promoter variants with mutations in the CDE leading to a decreased interaction with E2F show a delayed derepression. In contrast, the G4 mutant, which exhibits increased E2F binding, mainly due to gaining the ability to bind E2F-4, is upregulated prematurely. These results clearly suggest that E2F binding influences the timing of derepression. (iii) Previous studies have shown that the overexpression of E2F-1 leads to an up-regulation of the cyclin A gene (Ref.6, Ref.35). It is likely that this effect of E2F-1 overexpression can be explained in the same way as the effect of the G4 mutation.

The most plausible interpretation for these observations is that E2F acts as an anti-repressor on the interaction of CDF-1 with the CDE-CHR in the cyclin A promoter. This would offer an explanation (i) for the induction of the cyclin A gene by ectopic E2F expression (Ref.6, Ref.35), (ii) for the earlier derepression of the cyclin A promoter by increasing the affinity for E2F (mutant G4) and (iii) for the delay in up-regulation of promoter mutants showing decreased E2F binding (mutants C2A9 and C2G4A9). This model would also explain why the cdc25C promoter is upregulated later than the cyclin A gene (Ref.54): The cdc25C promoter is regulated by CDF-1 only, so that its derepression is solely dependent on its inactivation in S/G_2. In the case of the cyclin A gene, the same mechanism is operational, but competiton with E2F for DNA-binding would lead to an earlier dissociation at least of a fraction of CDF-1 from its cognate DNA binding site. Such a scenario is also supported by the fact that the derepression of the cyclin A gene directly follows the up-regulation of E2F in late G_1 (Ref.28).

The competition of E2F and CDF-1 can be investigated further by the provision of recombinant CDF-1, so that in vitro binding reactions can be carried out at a low ratio of probe:protein. The evidence obtained in the present study taken together with previous results strongly supports the idea of E2F acting as an anti-repressor at the cyclin A CDE-CHR motif.

The promoter of cyclin A, and derivatives of this promoter which retain the binding motifs for E2F and CDF-1 may be used, when operably linked to a structural gene, to provide constructs which are regulated in a cell cycle specific manner. These constructs may be used *in vitro* or *in vivo* in eukaryotic, particularly mammalian including rodent, primate (e.g. human) cells, to achieve cell cycle dependent expression of genes. Derivatives of the promoter are sequences which differ in regions away from the specific binding sites for E2F and CDF-1 by a degree which does not substantially affect the ability of the promoter to bind such factors. This can encompass very substantial changes to the native cyclin A promoter region, although promoters which have a homology of at least 60%, e.g. at least 70, 80, 90, 95 or 99% to a corresponding portion of the cyclin A promoter may be used.

Suitable genes which may be operably linked include genes encoding cytokines including interleukins such as IL-1, IL-2 isoforms and TNFα, HLA antigens or other cell suface antigens, receptors, cytotoxic or cytostatic proteins and enzymes such as those capable of converting a prodrug into an acitve drug such as tk, nitroreductase, β-glucuronidase, carboxypeptidase, cytosine deaminase or alkaline phosphatase.

The cyclin A promoter region may also comprise or be linked to further regulatory regions, for example regions which regulate expression in a tissue specific manner.

Materials and methods.

Cell culture, DNA transfection and luciferase assays.

NIH3T3 cells were cultured in Dulbecco-Vogt modified Eagle medium (DMEM) supplemented with 10% fetal calf serum, penicillin and streptomycin. HeLa cells were grown in DMEM plus 5% newborn calf serum. NIH3T3 cells were transfected by the DEAE dextran technique as described (Ref.24). For synchronization in G_0, cells were maintained in serum free medium for 2 days 12 hrs. after transfection and restimulated with 10% FCS for the time periods indicated in Figure 24. Determination of luciferase activities and standardization of results were performed as published (Ref.24).

Luciferase constructs.

The cdc25C and B-myb promoter-driven luciferase constructs have been described elsewhere (Ref.24, Ref.54, Ref.22). Mutations were introduced by PCR strategies as described (Ref.54). All PCR-amplified fragments were verified by DNA sequencing.

Fractionation of nuclear extract.

Nuclear extracts were prepared from HeLa suspension cultures in high salt extraction buffer (Ref.8) in the presence of the protease inhibitors leupeptin (50 ng/ml), pepstatin A (5 μg/ml), and aprotinin (80 ng/ml) and applied to a EDF-Sepharose (Merck, Darmstadt, Germany) column in a buffer containing 50 mM Tris-HCl , pH 7.5, 10% glycerol and 3mM DTT. Elution was performed by a gradient of 0.1 to 1M NaCl. CDF-1 eluted at a salt concentration of 150 mM, while all E2F complexes eluted at 300-400 mM.

Electrophoretic mobility shift assay (EMSA).

EMSA conditions and oligonucleotide probes have been published previously (Ref.22). The sequences of the cyclin A promoter mutants are indicated in Fig. 22A. The recombinant E2F and DP-1 proteins and their use in EMSAs has been described elsewhere (Ref.53). Antibodies specific for E2F-1 (KH-95) and E2F-3 (N-20) or E2F-4 (C-108) and used for supershifts were purchased from Santa-Cruz Biotechnology.

EXAMPLE 6: CDF-1 BINDING TO CDE/CHR IS REGULATED BY PHOSPHORYLATION AND DEPHOSPHORYLATION.

As indicated in Example 4 above, a potential mechanism uderlying the phase-specific DNA-binding of CDF-1 is protein modification by phosphorylation or dephosphorylation. To examine this in further detail, the effect of cyclin E-cdk2 on the DNA binding of CDF-1 was examined.

100 μl extracts of Sf9 cells containing baculovirus producing cyclin E and cdk2 were mixed with 5μl cyclin E antibody (Santa Cruze, HE11)n and 5μl protein A agarose (Pharmacia) and incubated for 1 hour on ice with occasional mixing. The agarose was washed twice in an NP-40 washing buffer (150 mM NaCl; 1% NP-40;50mM Tris.Cl pH8.0; 1 mM DTT; proteinase inhibitors (5μg/ml aprotinin, 5μg/ml leupeptin, 1μg/ml pepstatin and 0.2 mM PMSF); and phosphatase inhibitors (10 mM β-glycerophosphate, 1 mM NaF and 100 μM Na-ortho-vanadate).

The cyclin E-cdk2 complexes on agarose were then ready to be used for *in vitro* phosphorylation assays.

An extract made from G_0 cells was subjected to *in vitro* phosphorylation by baculovirus-produced cyclin E-cdk2 kinase, using 250 μg nuclear extract per 50 μg of the Sf9 extract. The kinase buffer used was 50mM Tris.Cl pH8.0; 10 mM $MgCl_2$; 50 μM ATP; 10% glycerol; 0.5 mM EDTA; 0.5 mM DTT; and phosphatase inhibitors as described above.

Using either the affinity purified cyclin E-cdk2 kinase or crude SF9 extract we found the DNA binding activity of CDF-1, as determined by EMSA as described above, was increased.

To prove that the phosphorylation of CDF-1 by cyclin E-cdk2 kinase can enhance binding, the experiment was repeated using nuclear extract from cells 16 hours post-stimulation (S/G_2 phase), in which CDF-1 binding activity is normally diminished. It was found that the treatment with cyclin E-

cdk2 kinase restored the binding activity.

To further confirm this effect, G_0 cells were treated instead with bacterial alkaline phosphatase (1μl BAP, Sigma, for 30 min at room temp per 100 mg extract). This markedly reduced the DNA binding activity.

Thus phosphorylation up-regulates and dephosphorylation down-regulates the DNA binding functions of CDF-1.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

References:

Ref.1 Bennett, J.D., *et al* (1996) Oncogene 13, 1073-1082.

Ref.2 Courey, A.J. and Tjian, R. (1988) Cell 55, 887-898.

Ref.3 Courtois, S.J., *et al* (1990) Nucl. Acids Res. 57-64.

Ref.4 Coustry, F., *et al* (1995) J. Biol. Chem. 270, 468-475.

Ref.5 Das, G. and Herr, W. (1995) Nature 374, 657-660.

Ref.6 DeGregori, J., *et al* (1995). Mol. Cell. Biol., 15, 215-4224.

Ref.7 Desbarats, L., *et al* (1996) Genes Dev. 10, 447-460.

Ref.8 Dignam, J.D., *et al* (1983). Nucl. Acids Res., 11, 1475-1489.

Ref.9 Good, L. and Nazar, R. (1992) Nucl. Acids Res. 20, 4934.

Ref.10 Goodrich, J.A. and Tjian, R. (1994) Curr. Opin. Cell Biol. 6, 403-409.

Ref.11 Helin, K. and Harlow, E. (1993) Trends Cell Biol. 3, 43-46.

Ref.12 Henglein, B., *et al* (1994). Proc. Natl. Acad. Sci. USA, 91, 5490-5494.

Ref.13 Hoey, T., *et al* (1993) Cell 72, 247-260.

Ref.14 Hsiao, K.M., *et al* (1994). Genes Dev., 8424,

1526-1537.

Ref.15 Huet, X., *et al* (1996). Mol. Cell. Biol., 16, 3789-3798.

Ref.16 Johnson, A.D. (1995) Cell 81, 655-658.

Ref.17 Johnson, D.G., *et al* (1994). Genes Dev., 8424, 1514-1525.

Ref.18 Klehr, D., *et al* (1991) Biochemistry 30, 1264-1270.

Ref.19 La Thangue, N.B. (1996) Biochem. Soc. Transactions 24, 54-59.

Ref.20 Lam, E.W. & Watson, R.J. (1993). EMBO J., 12, 2705-13.

Ref.21 Li, X.-Y., *et al* (1992) J. Biol. Chem. 267, 8984-8990.

Ref.22 Liu, N., *et al* (1996). Nucl. Acids Res., 24, 2905-2910.

Ref.23 Lowings, P., *et al* (1992) Mol. Cell. Biol. 12, 3653-62.

Ref.24 Lucibello, F.C., *et al* (1995). EMBO J., 14, 132-142.

Ref.25 Mantovani, R., *et al* (1992) EMBO J. 11, 3315-22.

Ref.26 Mermod, N., *et al* (1989) Cell 58, 741-753.

Ref.27 Mitchell, P.J. and Tjian, R. (1989) Science 245, 371-378.

Ref.28 Müller, R. (1995) Trends Genet., 11, 173-178.

Ref.29 Ohtani, K., *et al* (1996). Mol. Cell. Biol., 16, 6977-6984.

Ref.30 Pfeifer, G.P., *et al* (1989) Science 246, 810-813.

Ref.31 Regier, J.L., *et al* (1993) Proc. Natl. Acad. Sci. USA 90, 883-887.

Ref.32 Salminen, A., *et al* (1991) J. Cell Biol. 115, 905-917.

Ref.33 Sardet, C., *et al* (1995). Proc. Natl. Acad. Sci. USA, 92, 2403-2407.

Ref.34 Schreier, T., *et al* (1990) J. Biol. Chem. 265, 21247-21253.

Ref.35 Schulze, A., *et al* (1995). Proc. Natl. Acad. Sci. U.S.A., 92, 11264-11268.

Ref.36 Seipel, K., *et al* (1992) EMBO J. 13, 4961-4968.

Ref.37 Shibata, K., *et al* (1992) J. Biol. Chem. 267,

20584-8.

Ref.38 Sinha, S., *et al* (1995) J. Biol. Chem. 92, 1624-1628.

Ref.39 Smale, S.T. and Baltimore, D. (1989) Cell 57, 103-13.

Ref.40 Tommasi, S. & Pfeifer, G.P. (1995). Mol. Cell. Biol., 15, 6901-6913.

Ref.41 Urban, M.B. and Baeuerle, P.A. (1990) Genes Dev. 4, 1975-1984.

Ref.42 van Hujisduijnen, H.R., *et al* (1990) EMBO J. 9, 3119-3127.

Ref.43 Webster, N.J.G., *et al* (1989) EMBO J. 8, 1441-1446.

Ref.44 Weinberg, R.A. (1995) Cell 81, 323-330.

Ref.45 Weintraub, S.J., *et al* (1995) Nature 375, 812-815.

Ref.46 Wilson, R.E., *et al* (1989) Cancer Res 49, 711-6.

Ref.47 Yaffe, D. and Saxel, O. (1977) Nature (Lond.) 270, 725-727.

Ref.48 Zenke, M., *et al* (1986) EMBO J. 5, 387-397.

Ref.49 Zerfass-Thome, K., *et al* (1997). Mol. Cell. Biol., 17, 407-415.

Ref.50 Zhu, L., *et al* (1995). Mol. Cell. Biol., 15, 3552-3562.

Ref.51 Zwicker, J. & Müller, R. (1997). Trends Genet., 13, 3-6.

Ref.52 Zwicker, J., *et al* (1995) Nucleic Acids Res. 23, 3822-3830.

Ref.53 Zwicker, J., *et al* (1996) Science 271, 1595-1597.

Ref.54 Zwicker, J., *et al* (1995) EMBO J. 14, 4514-4522.

Although the foregoing invention has been described in some detail by way of illustration and example for the purposes of clarity and understanding, it will be readily apparent to those of skill in the art that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

CLAIMS

1. An assay for a modulator of viral replication, wherein said replication is facilitated by suppression of promoter repression mediated by CDF-1 due to the action of a viral protein, which assay comprises:

a) bringing said viral protein into contact with one or both of:

(i) a CDF-1 protein; and

(ii) a DNA sequence comprising a CDF-1 binding site,

in the presence of a putative modulator under conditions wherein in the absence of modulator the viral protein is able to induce inhibition of CDF-1 binding to DNA; and

b) measuring the degree of modulation of binding of CDF-1 to DNA.

2. An assay according to claim 1 which is performed in mammalian cells in which said viral protein is expressed.

3. An assay according to claim 2 wherein said cells are in synchronous culture.

4. An assay according to any one of the preceding claims wherein said viral protein is selected from the group of a human papilloma virus E7 protein, an E7 mutant and SV40 large T antigen.

5. An assay according to any one of claims 1 to 3 wherein the CDF-1 binding site comprises SEQ ID NO:1, the binding site being operably linked to a reporter gene construct.

6. An assay according to claim 5 wherein said binding site is a cdc25 promoter.

7. An assay according to claim 5 or 6 wherein said reporter gene is luciferase, chloramphenicol acetyl transferase, β-galactosidase, or green fluroescent protein.

8. An assay according to any one of claims 1 to 3 wherein the measuring is carried out by an electrophoretic mobility shift assay (EMSA) to determine the binding of CDF-1 to a CDF-1 binding site.

9. An assay according to any one of claims 1 to 3 in the form of a two-hybrid assay.

10. A purified CDF-1 protein preparation wherein said CDF-1 is enriched at least 10,000 fold over cellular components with which it is naturally associated.

11. An assay for a modulator of cell cycle progression which comprises providing

(i) a cell in which CDF-1 represses transcription of at least one gene by binding to a CDE-CHR binding site; and

(ii) a putative modulator; and

determining the effect of said modulator on the phosphorylation of CDF-1.

12. An assay for a modulator of cell cycle progression which comprises providing

(i) CDF-1 in a phosphorylated form; and

(ii) a putative modulator,

under reaction conditions in which CDF-1 is, in the absence of the modulator, subjected to a dephosphorylation reaction; and determining the effect of the modulator on the dephosphorylation reaction.

13. An assay according to claim 12 wherein the reaction conditions comprise supplying a phosphatase.

14. An assay for a modulator of cell cycle progression which comprises providing

(i) CDF-1 in a dephosphorylated form; and

(ii) a putative modulator,

under reaction conditions in which CDF-1 is, in the absence of the modulator, subjected to a phosphorylation reaction; and

determining the effect of the modulator on the phosphorylation reaction.

15. An assay according to claim 14 wherein the reaction conditions comprise supplying a kinase.

16. An assay according to any one of claims 12 to 15 wherein the phosphorylation status of CDF-1 is measured by EMSA.

17. The purified preparation of CDF-1 protein of claim 11 obtainable by:

a) providing a nuclear extract of mammalian cells;
b) precipitating said extract in the presence of 30% ammonium sulphate;
c) dialysing said precipitate at pH7.5 in the presence of 60mM NaCl to resuspend a CDF-1 fraction;
d) passing said fraction through a first DEAE anion exchange colum using an NaCl gradient of 50-1000 mM and collecting the 100-130 mM fraction;
e) passing said fraction from (d) through an EDF anion exchange column using an NaCl gradient of 120-1000 mM and collecting the 120-180 mM fraction;
f) centrifuging the fraction of (e) through a 5-20% glycerol gradient and collecting the 5-7% fraction;
g) passing the fraction of (f) through a cation exchange column using an NaCl gradient of 100-1000 mM and collecting the fraction above 800mM;
h) purifying the fraction of (g) using sequence specific DNA affinity purification using a DNA capture probe comprising a CDF-1 binding site; and
i) recovering the CDF-1.

18. An isolated fragment of the CDF-1 which retains the ability to bind to the cdc25 promoter CDF-1 binding site.

19. An antibody capable of binding to CDF-1.

20. A cDNA encoding CDF-1, said cDNA being obtainable by

probing an expression library with the antibody of claim 17, and selecting a clone which expresses a polypeptide capable of binding said antibody.

21. A cDNA encoding CDF-1, said cDNA being obtainable by treating a expression library with a G_1 kinase, and selecting a cDNA encoding CDF-1 with a probe comprising a CDE-CHR binding site.

22. A nucleic acid which has at least 70% homology to the cDNA of claim 20 or 21.

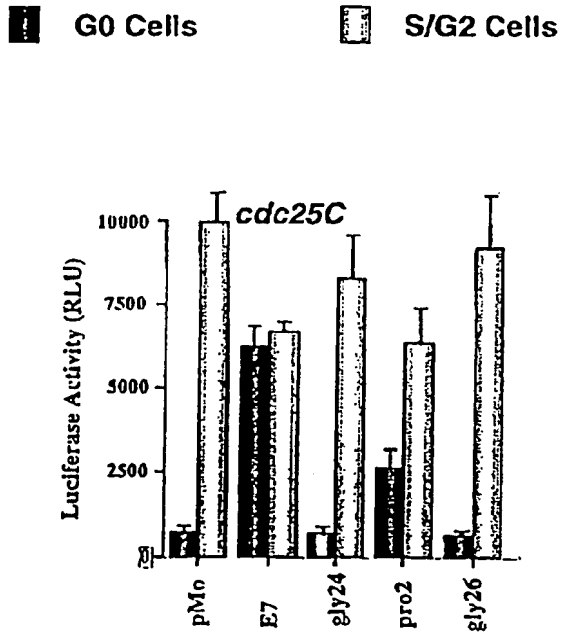

G0 Cells
S/G2 Cells
cdc25C
Luciferase Activity (RLU)
10000
7500
5000
2500
0
pMo
E7
gly24
pro2
gly26

Fig. 1

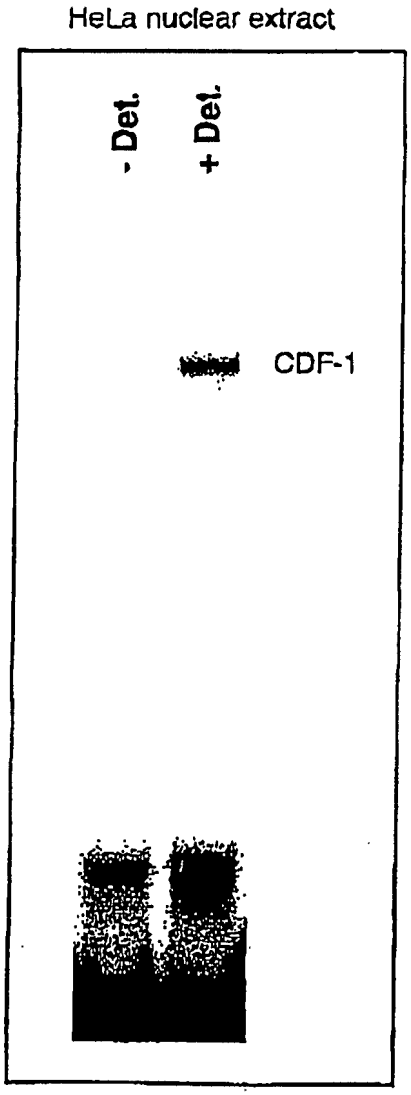


FIG. 2B

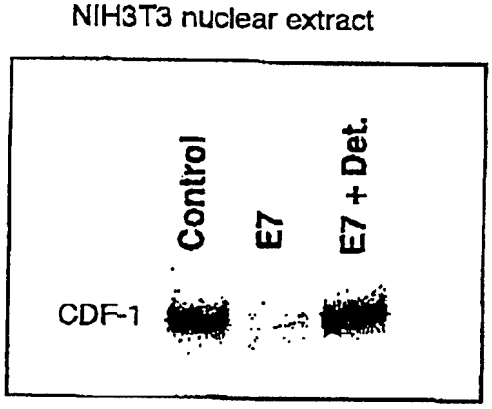


FIG. 2A

Fig. 2

BIOTIN-TTACTGGCTGGCGGAAGGTTTGAATGGTCAATTACTGGGCTGGCGGAAGGTTTGAATGGTCAA
TGACCCGACCGCCTTCCAAACTTACCAGTTAATGACCCGACCGCCTTCCAAACTTACCAGTT

Fig. 3:

Fig. 3

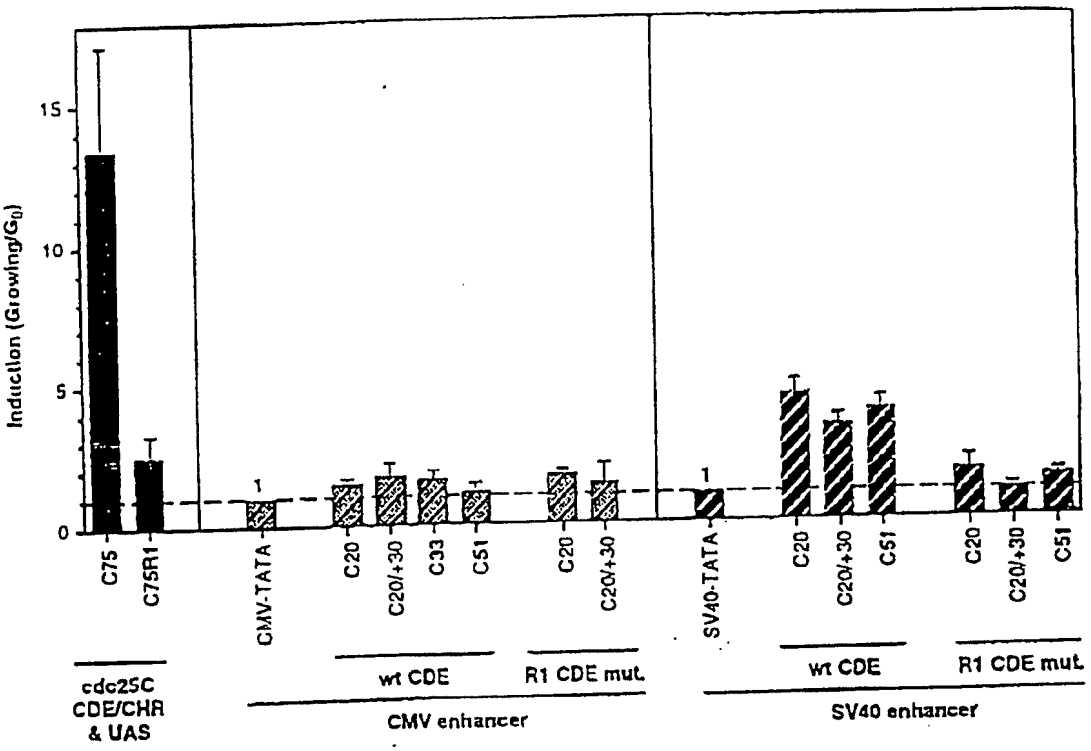

Induction (Growing/G0)
15
10
5
0
C75
C75R1
CMV-TATA
C20
C20/+30
C33
C51
C20
C20/+30
SV40-TATA
C20
C20/+30
C51
C20
C20/+30
C51
wt CDE
R1 CDE mut.
wt CDE
R1 CDE mut.
cdc25C CDE/CHR & UAS
CMV enhancer
SV40 enhancer

Fig. 4

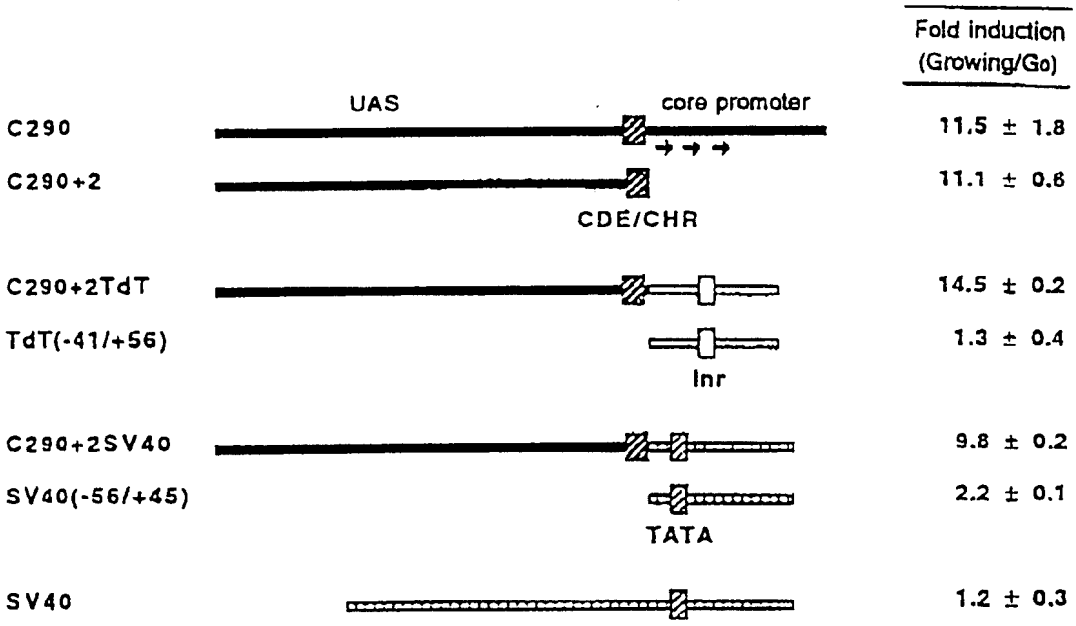
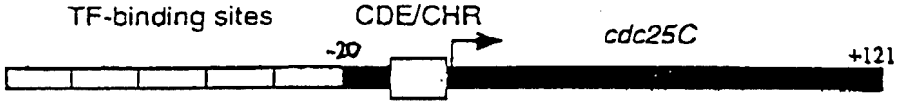
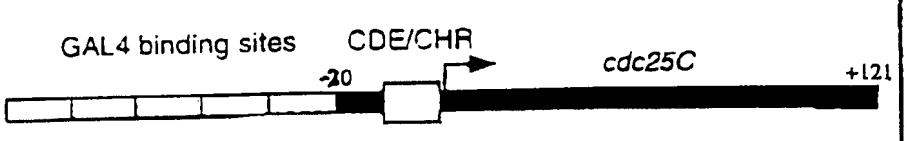

Fold induction (Growing/Go)
UAS
core promoter
C290
11.5 ± 1.8
C290+2
11.1 ± 0.6
CDE/CHR
C290+2TdT
14.5 ± 0.2
TdT(-41/+56)
1.3 ± 0.4
Inr
C290+2SV40
9.8 ± 0.2
SV40(-56/+45)
2.2 ± 0.1
TATA
SV40
1.2 ± 0.3

Fig. 5



Binding site	RLUs Growing / RLUs G_0		Cell cycle regulation
	Intact CDE/CHR	Mutated CDE/CHR	
NF-Y	5.8 ± 1.1	1.9 ± 0.6	3.0
NFκ-B	1.5 ± 0.4	1.4 ± 0.5	1.1
NF-I	1.9 ± 0.3	1.7 ± 0.1	1.1
Sp1	2.8 ± 0.2	1.3 ± 0.4	2.2

Fig. 6



Type	Activation domain	CDE/CHR-mediated cell cycle regulation	
Gln-rich	Oct-N	2.6 ± 0.4	+
	Sp1-Q1	2.1 ± 0.3	±
	NF-Y (GAYA-6)	2.8 ± 0.2	+
Ser-, Thr-, Gln-rich	NF-Y (GAYA-5)	3.5 ± 0.5	+
	NF-Y (GAYA-11)	5.5 ± 1.5	++
Ser-, Thr-rich	ITF2	1.4 ± 0.3	-
Acidic	VP-16	1.1 ± 0.3	-
	Myc	1.5 ± 0.3	-
Pro-rich	CTF	1.3 ± 0.2	-
-	none (vector control)	1.5 ± 0.3	-

Fig. 7

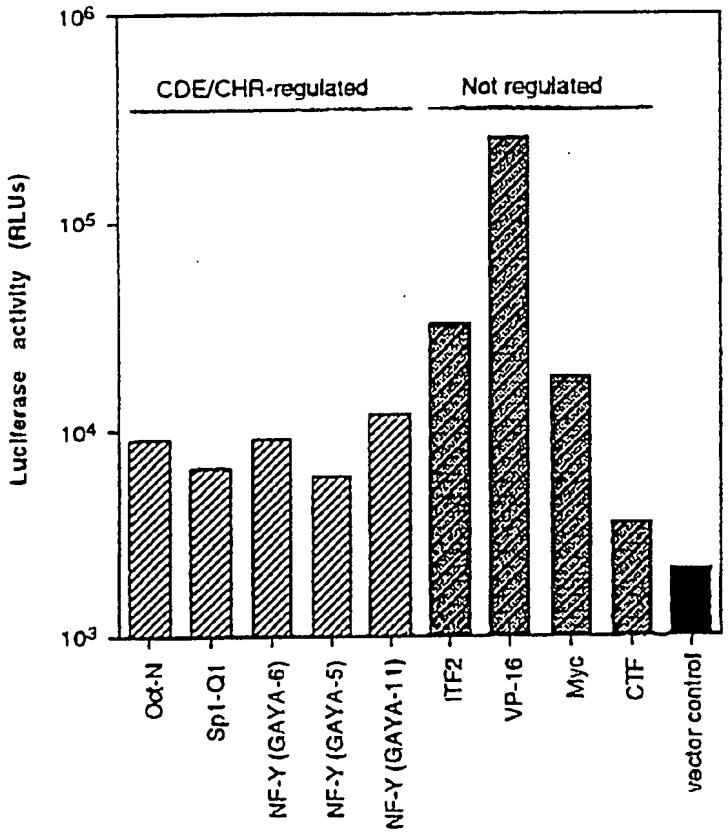

Luciferase activity (RLUs)
10^6
10^5
10^4
10^3
CDE/CHR-regulated
Not regulated
Oct-N
Sp1-Q1
NF-Y (GAYA-6)
NF-Y (GAYA-5)
NF-Y (GAYA-11)
ITF2
VP-16
Myc
CTF
vector control

Fig. 8

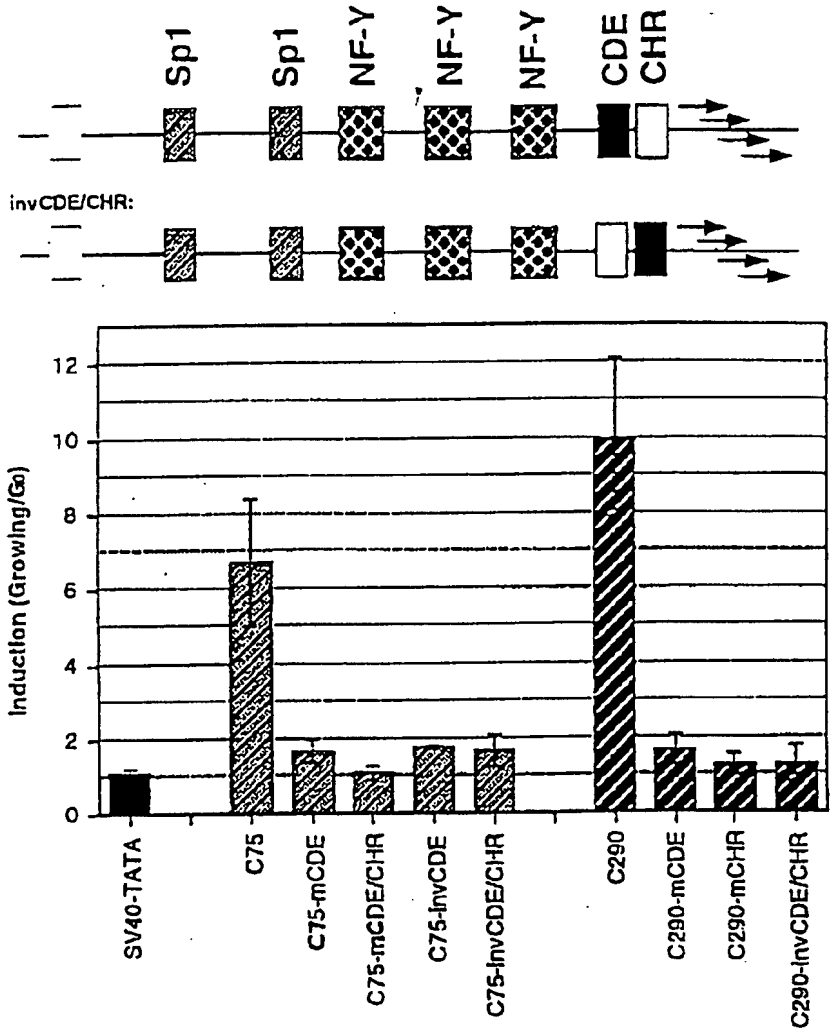

Sp1
Sp1
NF-Y
NF-Y
NF-Y
CDE
CHR
invCDE/CHR:
Induction (Growing/Go)
0
2
4
6
8
10
12
SV40-TATA
C75
C75-mCDE
C75-mCDE/CHR
C75-invCDE
C75-invCDE/CHR
C290
C290-mCDE
C290-mCHR
C290-invCDE/CHR

Fig. 9

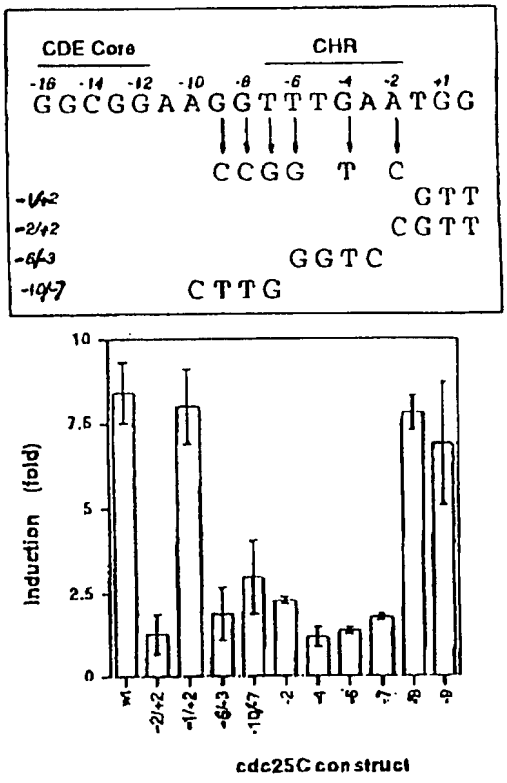

CDE Core
CHR
-16 -14 -12 -10 -8 -6 -4 -2 +1
GGCGGAAGGTTTGAATGG
CCGG T C
GTT
CGTT
GGTC
CTTG
Induction (fold)
10
7.5
5
2.5
0
cdc25C construct

Fig. 10

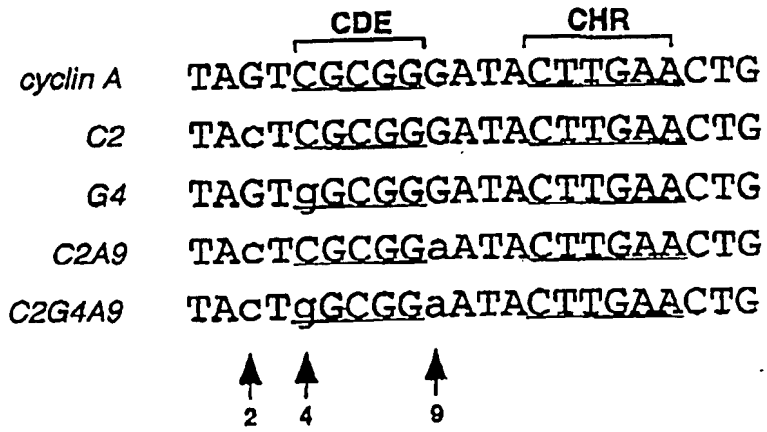

CDE
CHR
cyclin A TAGTCGCGGGATACTTGAACTG
C2 TAcTCGCGGGATACTTGAACTG
G4 TAGTgGCGGGATACTTGAACTG
C2A9 TAcTCGCGGaATACTTGAACTG
C2G4A9 TAcTgGCGGaATACTTGAACTG
2
4
9

Fig. 11

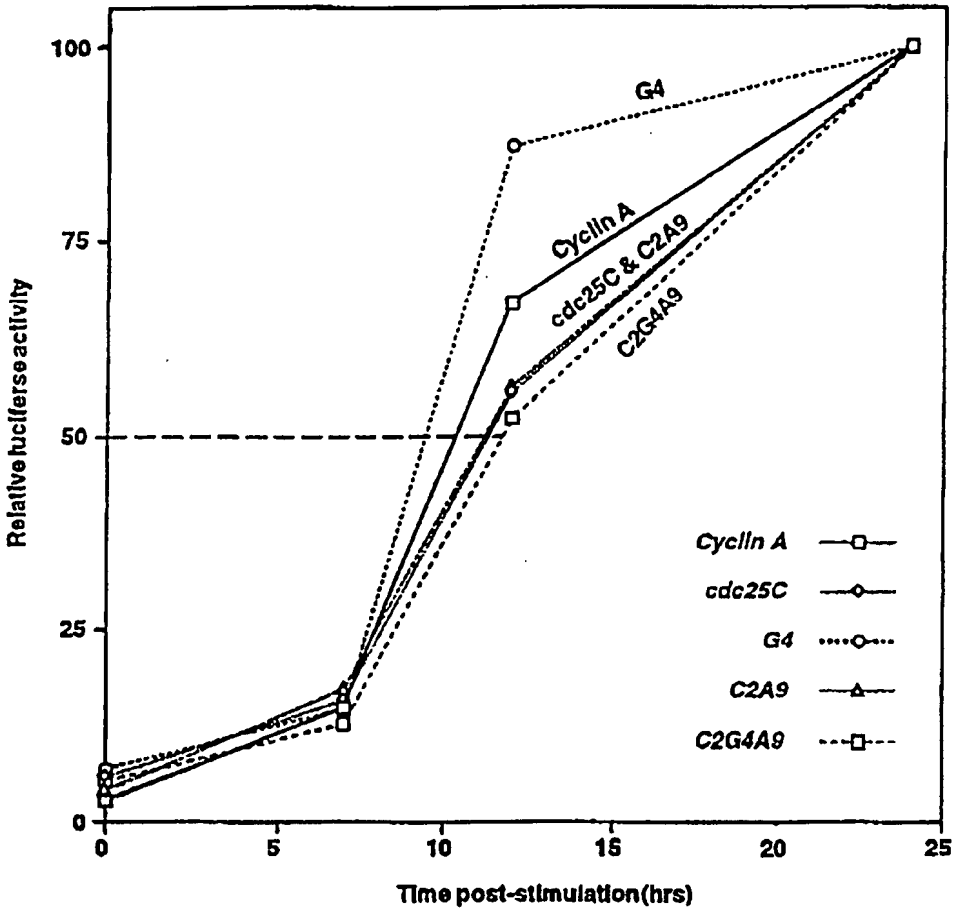

100
75
50
25
0
0
5
10
15
20
25
Relative luciferse activity
Time post-stimulation(hrs)
G4
Cyclin A
cdc25C & C2A9
C2G4A9
Cyclin A
cdc25C
G4
C2A9
C2G4A9

Fig. 12

INTERNATIONAL SEARCH REPORT

Interr. ial Application No
PCT/GB 98/03486

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 C12Q1/68 G01N33/68 C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 06943 A (PROLIFIX LTD.; BEHRINGWERKE AG (DE); MÜLLER ROLF) 7 March 1996 cited in the application	9,16-21
A	see page 10, line 10 - page 13, line 14 see page 36; claims 12-15	10,11,13
P,X, L	EP 0 859 008 A (HOECHST AG (DE); MÜLLER ROLF; LIU N.; ZWICKER J.; SEDLACEK H.-H.) 19 August 1998 L: priority see page 3, line 21 - page 5, line 7 see page 14, line 48-55 see page 15, line 45 - page 16, line 33 see page 21; claims 19-24	9,16

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 February 1999	09/03/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Macchia, G

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Interr. ıal Application No
PCT/GB 98/03486

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X, L	EP 0 860 445 A (HOECHST AG (DE); MÜLLER ROLF; LIU N.; ZWICKER J.; SEDLACEK H.-H.) 26 August 1998 see page 3, line 18 - page 5, line 23 see page 19, line 42-49 see page 20, line 37 - page 21, line 23 see page 31; claims 22,23	9,16
P,X	LIU N. ET AL.: "CDF-1, a novel E2F-unrelated factor, interacts with cycle-regulated repressor elements in multiple promoters" NUCLEIC ACIDS RESEARCH, vol. 25, no. 24, 15 December 1997, pages 4915-4920, XP002093649 see abstract see page 4916, left-hand column, paragraph 5 - right-hand column, paragraph 1 see page 4917, right-hand column, paragraph 1	9,16

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern al Application No
PCT/GB 98/03486

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9606943 A	07-03-1996	AU 3263995 A	22-03-1996
		AU 690733 B	30-04-1998
		AU 3387495 A	22-03-1996
		AU 696593 B	17-09-1998
		AU 3473295 A	22-03-1996
		AU 690336 B	23-04-1998
		AU 3473395 A	22-03-1996
		AU 695458 B	13-08-1998
		AU 3518595 A	22-03-1996
		CA 2198080 A	07-03-1996
		CA 2198462 A	07-03-1996
		CA 2198472 A	07-03-1996
		CA 2198473 A	07-03-1996
		CA 2198474 A	07-03-1996
		CN 1158144 A	27-08-1997
		WO 9606938 A	07-03-1996
		WO 9606939 A	07-03-1996
		WO 9606940 A	07-03-1996
		WO 9606941 A	07-03-1996
		EP 0802985 A	29-10-1997
		EP 0777739 A	11-06-1997
		EP 0804601 A	05-11-1997
		EP 0807183 A	19-11-1997
		EP 0777740 A	11-06-1997
		FI 970674 A	21-04-1997
		FI 970768 A	24-02-1997
		JP 10507626 T	28-07-1998
		JP 10505234 T	26-05-1998
		JP 10506526 T	30-06-1998
		JP 10508188 T	18-08-1998
		US 5830880 A	03-11-1998
		NO 970756 A	26-02-1997
EP 0859008 A	19-08-1998	EP 0860445 A	26-08-1998
		AU 5539898 A	20-08-1998
		CA 2224123 A	18-08-1998
		CN 1197077 A	28-10-1998
		CZ 9800460 A	16-09-1998
		HU 9800329 A	28-09-1998
		JP 10229888 A	02-09-1998
		PL 324919 A	31-08-1998
EP 0860445 A	26-08-1998	AU 5539898 A	20-08-1998
		CA 2224123 A	18-08-1998
		CN 1197077 A	28-10-1998
		CZ 9800460 A	16-09-1998
		EP 0859008 A	19-08-1998
		HU 9800329 A	28-09-1998
		JP 10229888 A	02-09-1998
		PL 324919 A	31-08-1998